LAMAR UNIVERSITY-BEAUMONT
0 00 01 1106992 8

DISCARDED

LIBRARY OF
LAMAR STATE COLLEGE OF TECHNOLOGY

PROCEEDINGS

OF THE

ROYAL SOCIETY OF LONDON

SERIES A

CONTAINING PAPERS OF A MATHEMATICAL AND PHYSICAL CHARACTER.

VOL. CXV

LONDON:

PRINTED FOR THE ROYAL SOCIETY AND SOLD BY
HARRISON AND SONS, LTD., ST. MARTIN'S LANE,
PRINTERS IN ORDINARY TO HIS MAJESTY.

AUGUST, 1927.

LONDON
HARRISON AND SONS, LTD., PRINTERS IN ORDINARY TO HIS MAJESTY,
ST. MARTIN'S LANE.

CONTENTS.

SERIES A. VOL. CXV.

No. 770.—June 1, 1927.

PAGE

The Zeeman Effect and Spherical Harmonics. By C. G. Darwin, F.R.S. 1

Electron Emission under the Influence of Chemical Action at Higher Gas Pressures, and some Photo-electric Experiments with Liquid Alloys. By O. W. Richardson, F.R.S., and M. Brotherton 20

Gaseous Combustion at High Pressures. Part VII.—A Spectrographic Investigation of the Ultra-Violet Radiation from Carbonic Oxide-Oxygen (or -Air) Explosions. By W. A. Bone, F.R.S., and D. M. Newitt. (Plate 1) 41

An Investigation of the Rate of Growth of Crystals in Different Directions. By M. Bentivoglio. Communicated by Sir Henry Miers, F.R.S. 59

The Latent Heat of Vaporization of Benzene at Temperatures above the Boiling Point. By J. A. Sutcliffe, F. C. Lay and W. Ll. Prichard. Communicated by D. L. Chapman, F.R.S. 88

The Relations Connecting the Angle-Sums and Volume of a Polytope in Space of n Dimensions. By D. M. Y. Sommerville. Communicated by Major P. A. MacMahon, F.R.S. 103

Thermal Changes in Iron-Manganese Alloys, Low in Carbon. By Sir Robert Hadfield, F.R.S. 120

Tensile Tests on Alloy Crystals.—Part I. By C. F. Elam. Communicated by H. C. H. Carpenter, F.R.S. (Plate 2) 133

Tensile Tests on Alloy Crystals.—Part II. By C. F. Elam. Communicated by H. C. H. Carpenter, F.R.S. (Plates 3 and 4) 148

Tensile Tests on Alloy Crystals.—Part III. By C. F. Elam. Communicated by H. C. H. Carpenter, F.R.S. 167

Meteorological Perturbations of Tides and Currents in an Unlimited Channel Rotating with the Earth. By H. Horrocks. Communicated by Sir Joseph Larmor, F.R.S. 170

Generalised Sturm-Liouville Expansions in Series of Pairs of Related Functions. By H. Horrocks. Communicated by Sir Joseph Larmor, F.R.S. 184

The Expansion of Charcoal on Sorption of Carbon Dioxide. By F. T. Meehan. Communicated by Sir William Hardy, F.R.S. 199

PAGE

A Contribution to Modern Ideas on the Quantum Theory. By H. T. Flint and J. W. Fisher. Communicated by O. W. Richardson, F.R.S. 208

Homogeneous Reactions Involving Complex Molecules. The Kinetics of the Decomposition of Gaseous Dimethyl Ether. By C. N. Hinshelwood and P. J. Askey. Communicated by H. Hartley, F.R.S. 215

An Experimental Test of the Dipole Theory of Adsorption. By W. G. Palmer. Communicated by Sir William Pope, F.R.S. 227

The Thermal and Electrical Conductivity of a Single Crystal of Aluminium. By Ezer Griffiths, F.R.S. 236

No. 771.—July 1, 1927.

On Certain Average Characteristics of World-Wide Magnetic Disturbance. By S. Chapman, F.R.S. 242

The Method of Images in Some Problems of Surface Waves. By T. H. Havelock, F.R.S. 268

The Emission of Soft X-Rays by Different Elements. By O. W. Richardson, F.R.S., and F. S. Robertson 280

On the Nature of Wireless Signal Variations.—I. By E. V. Appleton, F.R.S., and J. A. Ratcliffe. (Plate 5) 291

On the Nature of Wireless Signal Variations.—II. By E. V. Appleton, F.R.S., and J. A. Ratcliffe 305

The Dissociation of Carbon Dioxide at High Temperatures. By R. W. Fenning and H. T. Tizard, F.R.S. 318

The Equation of State of a Gaseous Mixture. By J. E. Lennard-Jones and W. R. Cook. Communicated by R. H. Fowler, F.R.S. 334

The Influence of Boundary Films on Corrosive Action. By L. H. Callendar. Communicated by H. C. H. Carpenter, F.R.S. 349

The Band Spectrum of Water Vapour. By D. Jack. Communicated by O. W. Richardson, F.R.S. (Plate 6) 373

The Brownian Movement of a Galvanometer Coil and the Influence of the Temperature of the Outer Circuit. By L. S. Ornstein, H. C. Burger, J. Taylor and W. Clarkson. Communicated by O. W. Richardson, F.R.S. 391

On the Reflexion of Sound from a Porous Surface. By E. T. Paris. Communicated by F. E. Smith, F.R.S. 407

The Efficiency of K Series Emission by K Ionised Atoms. By L. H. Martin. Communicated by Sir Ernest Rutherford, P.R.S. 420

Currents Carried by Point-Discharges beneath Thunderclouds and Showers. By T. W. Wormell. Communicated by C. T. R. Wilson, F.R.S. (Plate 7) 443

The Crystal Structure of α-Manganese. By A. J. Bradley and J. Thewlis. Communicated by W. L. Bragg, F.R.S. 456

PAGE

The Rusting of Steel Surfaces in Contact. By G. A. Tomlinson. Communicated by Sir William Hardy, F.R.S. 472

A Note on the Specific Heat of the Hydrogen Molecule. By D. M. Dennison. Communicated by R. H. Fowler, F.R.S. 483

No. 772.—August 1, 1927.

Bakerian Lecture: A New Mass-Spectrograph and the Whole Number Rule. By F. W. Aston, F.R.S. (Plate 8) 487

On the Wave-Length of the Green Auroral Line in the Oxygen Spectrum. By J. C. McLennan, F.R.S., and J. H. McLeod. (Plate 9) 515

The Hydrogen Band Spectrum: New Band Systems in the Violet. By O. W. Richardson, F.R.S. 528

The Magnetic Anisotropy of Crystalline Nitrates and Carbonates. By K. S. Krishnan and C. V. Raman, F.R.S. 549

A New Differential Dilatometer for the Determination of Volume Changes during Solidification. By C. J. Smith. Communicated by C. H. Desch, F.R.S. 554

The Cause of the Colours shown during the Oxidation of Metallic Copper. By F. H. Constable. Communicated by T. M. Lowry, F.R.S. 570

On the Electric Moment of the Sulphur Complex. By A. M. Taylor and E. K. Rideal. Communicated by Sir Joseph J. Thomson, F.R.S. 589

The Total Ionisation due to the Absorption in Air of Slow Cathode Rays. By J. F. Lehmann and T. H. Osgood. Communicated by Sir Ernest Rutherford, P.R.S. 609

The Absorption of Slow Cathode Rays in Various Gases. By J. F. Lehmann. Communicated by Sir Ernest Rutherford, P.R.S. 624

A Simple Radioactive Method for the Photographic Measurement of the Integrated Intensity of X-Ray Spectra. By W. T. Astbury. Communicated by Sir William Bragg, F.R.S. 640

Further Developments of the Method of Obtaining Strong Magnetic Fields. By P. Kapitza. Communicated by Sir Ernest Rutherford, P.R.S. 658

The Combination of Nitrogen and Hydrogen Activated by Electrons. By A. Caress and E. K. Rideal. Communicated by T. M. Lowry, F.R.S. 684

A Contribution to the Mathematical Theory of Epidemics. By W. O. Kermack and A. G. McKendrick. Communicated by Sir Gilbert Walker, F.R.S. 700

Index 723

Minutes of Meetings: May 12, 19; June 2, 16, 23, 30.

PROCEEDINGS OF
THE ROYAL SOCIETY.

Section A.—Mathematical and Physical Sciences.

The Zeeman Effect and Spherical Harmonics.

By C. G. Darwin, F.R.S.

(Received March 23, 1927.)

1. The chief object of the present paper is to present a simple system of equations which are competent to determine the frequencies and intensities of the lines in the standard Zeeman effect. By the standard Zeeman effect is meant the type where the terms are given by Landé's "g" and "γ" formulæ, and where the multiplicity of the two sets of terms is the same. It will probably be held that the theory of such multiplets has been fairly completely understood for the last two years owing to the works of Sommerfeld, Heisenberg, Landé and Pauli,* and the main new contribution of the present work is that, whereas these writers gave formulæ valid only either in weak or strong field (except for doublets, where Heisenberg and Jordan† give the value for any strength), here we have complete formulæ from which the intensity of any component in any strength of field can be obtained merely from the solution of rather simple algebraic equations.

The proper attack on this problem would undoubtedly be by way of the recent work of Heisenberg‡ on the helium spectrum, and his still more recent work on complex spectra; but this theory is still in the making, so that it has not been practicable to apply it here, and to this extent our results are unverified. Their validity rests firstly on a complete verification of all known facts connected with both weak and strong fields (and intermediate fields for doublets), and also on a conformity with the general features of wave mechanics. The work of Heisenberg and Jordan could readily have been adapted to give all the results of the present paper; but it would have been harder to follow because the matrix methods are not so easy for most readers as are spherical harmonics.

If the wave mechanics are applied to a rotating body, in the formulation as originally given by Schrödinger, they lead to rotations which must have integers

* See for instance Andrade, 'The Structure of the Atom,' 3rd edition, Chapter XV.

† 'Z. f. Physik,' vol. 37, p. 263 (1926).

‡ 'Z. f. Physik,' vol. 39, p. 499 (1926); vol. 41, p. 239 (1927).

for their quantum numbers, and so are not directly applicable for the half quantum number that the spinning electron requires. If nevertheless we persist in the calculation of the problem of the spinning electron, we are led to a system of equations which exactly corresponds to the standard triplet. This fact furnished the starting point of the present investigation. A rotating body moving in an orbit can be made to give the complete system of any odd multiplicity.

For even multiplicities this is not so, since they seem to require solutions of the wave equation which are inadmissible if the characteristic functions ψ are to be one-valued. This is certainly a severe difficulty; a tentative way of meeting it has been suggested by Heisenberg, London,* and Dennison,† in proposing that $|\psi^2|$ is the quantity that should be one-valued. This certainly permits of what we may call half-harmonics for the rotating body, but it may be doubted if the matter ends there, for we have no assigned reason why these half-harmonics should not also apply for the revolution of the electron about the nucleus, and so such a rule at any rate wants safeguarding in order to ensure that the orbital motion shall always be of whole harmonic type, for otherwise the Balmer series would have members in $R/(n+\frac{1}{2})^2$. In fact the proposed modification implies that Schrödinger's principle is inadequate to decide between two sets of alternatives as the permissible states of a mechanical system. It is not our intention to enter into this matter, and we can deal with even multiplicities without doing so to a considerable extent, for when the equations as determined for odd multiplicities are set down it appears that they are also applicable in every respect for even, the only change being not from whole to half numbers, but merely from even numbers to odd.‡

We shall therefore give a development of the theory of an avowedly rather crude model, taking as its justification the fact that it gives all the observed results. The development of the work goes in the form of tesseral spherical harmonics, and this is an essentially anisotropic procedure. The consequence is that one of the laws of spectra—the law of combination of the quantum number j—is rather concealed. In § 7 we sketch a different development which brings out the meaning of j with full force.

* 'Z. f. Physik,' vol. 40, p. 193 (1926), footnote on p. 209.

† 'Nature,' vol. 119, p. 316 (1927).

‡ [*Added April*, 1927.—Since writing the paper I have had the advantage of discussing "half-harmonics" with Dr. Dennison. By making an insignificant alteration in the definition of the function $P_r^{s,t}$ (incorporated in the text) the present work is directly applicable to even multiplicities, when r, s and t are simultaneously all half-numbers. But the questions raised above remain outstanding and the method of § 7 is still not available.]

2. Our model consists of a charged spinning spherical body moving in a central field of force. Let $-e$, m, be the charge and mass, x, y, z the position of the centre, and $V(r)$ the potential energy at distance r. Then the motion of revolution contributes terms

$$\tfrac{1}{2}m(\dot{x}^2+\dot{y}^2+z^2)+eV(r)$$

to the Lagrangian function. Next take a set of Eulerian angles χ, λ, μ, and let the moment of inertia be I. This gives a contribution

$$\tfrac{1}{2}I(\dot{\chi}^2+\dot{\lambda}^2+\dot{\mu}+2\dot{\lambda}\dot{\mu}\cos\chi)=\tfrac{1}{2}I(\omega_x^2+\omega_y^2+\omega_z^2),$$

where ω_x, ω_y, ω_z are the components of spin about x, y, z. For an isotropic body it is possible to drop one of the three co-ordinates (μ), but perhaps the most convincing way of seeing this is to carry through the work without doing so, and to show that in fact it has no effect. Next we have the contribution arising from the magnetic field H along z. The orbital motion gives a contribution $-H\frac{e}{2c}(x\dot{y}-\dot{x}y)$. The spin gives $-HI\frac{e}{mc}\omega_z$ when we take e/mc for the ratio of magnetic moment to mechanical momentum, the hypothesis of Uhlenbeck and Goudsmit. Lastly, we have the interaction of the spin and the motion. The electric force has components $-V'(x/r)$, etc., so that according to Uhlenbeck and Goudsmit we should take a term

$$-I\frac{e}{mc^2}\frac{V'}{r}\{\omega_x(y\dot{z}-\dot{y}z)+\omega_y(z\dot{x}-\dot{z}x)+\omega_z(x\dot{y}-\dot{x}y)\}.$$

This overlooks Thomas's correction, but as we do not intend to study the absolute value of the multiplet separation, we shall simply take the outside factor as $-Im\,U(r)$.

We take all these terms together and convert into Hamiltonian form, neglecting the squares of the last two terms. The result, expressed in polar co-ordinates r, θ, ϕ, is

$$\begin{aligned}
&\frac{1}{2m}\left(P_r^2+\frac{1}{r^2}P_\theta^2+\frac{1}{r^2\sin^2\theta}P_\phi^2\right)-eV\\
&+\frac{1}{2I}\left\{P_\chi^2+\frac{1}{\sin^2\chi}(P_\lambda^2+P_\mu^2-2P_\lambda P_\mu\cos\chi)\right\}\\
&+\frac{eH}{2mc}(P_\phi+2P_\lambda)\\
&+U(r)\left\{(-P_\theta\sin\phi-P_\phi\cot\theta\cos\phi)\left(-P_\chi\sin\lambda+\frac{P_\mu-P_\lambda\cos\chi}{\sin\chi}\cos\lambda\right)\right.\\
&\left.+(P_\theta\cos\phi-P_\phi\cot\theta\sin\phi)\left(P_\chi\cos\lambda+\frac{P_\mu-P_\lambda\cos\chi}{\sin\chi}\sin\lambda\right)+P_\phi P_\lambda\right\}
\end{aligned}\tag{2.1}$$

The next stage is to form the wave equation. In doing so it is not of course sufficient just to write $P_\theta = \frac{h}{2\pi i}\frac{\partial}{\partial\theta}$, etc., but the variational method of Schrödinger provides an unambiguous answer. The method is not strictly applicable for the linear terms in H, and for these we do write $P_\phi = \frac{h}{2\pi i}\frac{\partial}{\partial\phi}$, etc. (it has been shown by Epstein* how the external magnetic field can be replaced by an electric current so as to make them quadratic). The result is

$$\begin{aligned}
W\psi = &-\frac{h^2}{8\pi^2 m}\left\{\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}+\frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right\}\psi - eV\psi \\
&-\frac{h^2}{8\pi^2 I}\left\{\frac{1}{\sin\chi}\frac{\partial}{\partial\chi}\sin\chi\frac{\partial}{\partial\chi}+\frac{1}{\sin^2\chi}\left(\frac{\partial^2}{\partial\lambda^2}+\frac{\partial^2}{\partial\mu^2}-2\cos\chi\frac{\partial^2}{\partial\lambda\,\partial\mu}\right)\right\}\psi \\
&+\frac{eHh}{4\pi mci}\left(\frac{\partial}{\partial\phi}+2\frac{\partial}{\partial\lambda}\right)\psi \\
&-\frac{h^2}{8\pi^2}U(r)\left\{e^{i(\phi-\lambda)}\left(\frac{\partial}{\partial\theta}+i\cot\theta\frac{\partial}{\partial\phi}\right)\left(\frac{\partial}{\partial\chi}+\frac{i}{\sin\chi}\left[\frac{\partial}{\partial\mu}-\cos\chi\frac{\partial}{\partial\lambda}\right]\right)\right. \\
&\left.+e^{-i(\phi-\lambda)}\left(\frac{\partial}{\partial\theta}-i\cot\theta\frac{\partial}{\partial\phi}\right)\left(\frac{\partial}{\partial\chi}-\frac{i}{\sin\chi}\left[\frac{\partial}{\partial\mu}-\cos\chi\frac{\partial}{\partial\lambda}\right]\right)+2\frac{\partial^2}{\partial\phi\,\partial\lambda}\right\}\psi.
\end{aligned} \tag{2.2}$$

3. We have to find the characteristic values and functions for this equation, and must start with the nul approximation which disregards the small terms. The orbital motion is solved by

$$W = W_{nk}, \qquad \psi = f_{nk}(r)\,P_k^u(\cos\theta)\,e^{iu\phi} \tag{3.1}$$

—the degeneracy in k that occurs in the hydrogen spectrum will not arise unless $V(r)$ is of the form $1/r$. The last two factors are the ordinary tesseral spherical harmonics, but it is rather convenient to make a slightly unusual definition for P_k^u; it is one which could advantageously be introduced generally. We define

$$P_k^u(z) = (k-u)!\,(1-z^2)^{\frac{u}{2}}\left(\frac{d}{dz}\right)^{k+u}\frac{(x^2-1)^k}{2^k k!}. \tag{3.2}$$

This is the usual form multiplied by $(k-u)!$ but the definition holds for negative values of u as well as positive, and the values are symmetrical about u zero. u admits of all integral values between $-k$ and k inclusive. The advantage of the introduction of the extra factor $(k-u)!$ is that all the scales of relation, etc., hold valid running right through $u=0$, which is not

* 'Proc. Nat. Acad.,' vol. 12, p. 634 (1926).

the case if it is omitted. We have not normalised the functions; indeed it is hardly ever an advantage to do so except in general theorems.

The following easily proved relations will be used later.

$$\left.\begin{aligned}\left(\frac{\partial}{\partial\theta} - i\cot\theta\,\frac{\partial}{\partial\phi}\right)\mathrm{P}_k^u(\cos\theta)\,e^{iu\phi} &= (k+u)\,\mathrm{P}_k^{u-1}e^{iu\phi}\\ \left(\frac{\partial}{\partial\theta} + i\cot\theta\,\frac{\partial}{\partial\phi}\right)\mathrm{P}_k^u(\cos\theta)\,e^{iu\phi} &= -(k-u)\,\mathrm{P}_k^{u+1}e^{iu\phi}\end{aligned}\right\} \tag{3.3}$$

Note that these relations automatically bar expressions like P_k^{k+1} and P_k^{-k-1}. We also have the normalizing relation

$$\int_0^{\pi}\int_0^{2\pi}\sin\theta\,d\theta\,d\phi\,|\,\mathrm{P}_k^u(\cos\theta)\,e^{iu\phi}\,|^2 = \frac{4\pi}{2k+1}(k+u)\,!\,(k-u)\,! \tag{3.4}$$

and the orthogonal relation

$$\int_0^{\pi}\int_0^{2\pi}\sin\theta\,d\theta\,d\phi\,\mathrm{P}_k^u(\cos\theta)\,e^{iu\phi}\,\mathrm{P}_{k'}^{u'}(\cos\theta)\,e^{-iu'\phi} = 0$$

unless both $k' = k$ and $u' = u$. Further if X stands for either $\sin\theta e^{-i\phi}$, $\cos\theta$ or $\sin\theta e^{i\phi}$

$$\int_0^{\pi}\int_0^{2\pi}\sin\theta\,d\theta\,d\phi\,\mathrm{P}_k^u e^{iu\phi}\,\mathrm{P}_{k'}^{u'}e^{-iu'\phi}\,.\,\mathrm{X}$$

vanishes always unless $k' = k \pm 1$. If $k' = k-1$, the only integrals of this type that do not vanish are

$$\left.\begin{aligned}\int_0^{\pi}\int_0^{2\pi}\sin\theta\,d\theta\,d\phi\,\mathrm{P}_k^u e^{iu\phi}\cdot\mathrm{P}_{k-1}^{u-1}e^{-i(u-1)\phi}\cdot\sin\theta e^{-i\phi} &= \frac{4\pi\,(k+u)!\,(k-u)!}{(2k+1)\,(2k-1)}\\ \int_0^{\pi}\int_0^{2\pi}\sin\theta\,d\theta\,d\phi\,\mathrm{P}_k^u e^{iu\phi}\,\mathrm{P}_{k-1}^u e^{-iu\phi}\cos\theta &= \frac{4\pi\,(k+u)\,!\,(k-u)\,!}{(2k+1)\,(2k-1)}\\ \int_0^{\pi}\int_0^{2\pi}\sin\theta\,d\theta\,d\phi\,\mathrm{P}_k^u e^{iu\phi}\,\mathrm{P}_{k-1}^{u+1}\,e^{-(u+1)\phi}\sin\theta e^{i\phi} &= -\frac{4\pi\,(k+u)\,!\,(k-u)!}{(2k+1)\,(2k-1)}\end{aligned}\right\} \tag{3.5}$$

The wave equation for a rotating body with two equal moments has been solved by Reiche* and shown to depend on hypergeometric functions. We here have a specially simple case on account of the equality of all the moments, but the solution is practically the same. We give it in a form which shows the close resemblance of the characteristic functions to spherical harmonics. The characteristic values are $\mathrm{W} = \dfrac{h^2}{8\pi^2\mathrm{I}}\,r\,(r+1)$ with $r = 0,\ 1,\ 2.\ .\ .\ .$, and the

* 'Z. f. Physik,' vol. 39, p. 444 (1926),

associated functions are $P_r^{s,t}(\cos\chi)\, e^{i(s\lambda+t\mu)}$ where $P_r^{s,t}(z)$ satisfies the equation

$$\frac{d}{dz}(1-z^2)\frac{dP}{dz}+\left[r(r+1)-\frac{s_2+t^2-2stz}{1-z^2}\right]P=0.$$

We take as the solution,

$$P_r^{s,t}(z)=(r-s)!\,(1-z)^{(s-t)/2}(1+z)^{(s+t)/2}\left(\frac{d}{dz}\right)^{r+s}(1-z)^{r+t}(1+z)^{r-t}. \quad (3.6)$$

Both s and t admit of all integral values between $-r$ and r inclusive, and the definition applies for both positive and negative values.*

If $t=0$, $P_r^{s,0}=(-2)^r r!\,P_r^s$. The following relations are easily proved

$$\left[\frac{\partial}{\partial\chi}+\frac{i}{\sin\chi}\left(\frac{\partial}{\partial\mu}-\cos\chi\frac{\partial}{\partial\lambda}\right)\right]P_r^{s,t}(\cos\chi)\,e^{i(s\lambda+t\mu)}=(r+s)\,P_r^{s-1,t}(\cos\chi)\,e^{i(s\lambda+t\mu)},$$

$$\left[\frac{\partial}{\partial\chi}-\frac{i}{\sin\chi}\left(\frac{\partial}{\partial\mu}-\cos\chi\frac{\partial}{\partial\lambda}\right)\right]P_r^{s,t}(\cos\chi)\,e^{i(s\lambda+t\mu)}=-(r-s)P_r^{s+1,t}(\cos\chi)e^{i(s\lambda+t\mu)} \quad (3.7)$$

We shall also require the normalising relation

$$\int_0^\pi\int_0^{2\pi}\int_0^{2\pi}\sin\chi\,d\chi\,d\lambda\,d\mu\,|\,P_r^{s,t}(\cos\chi)\,e^{i(s\lambda+t\mu)}\,|^2$$
$$=\frac{2^{2r+3}\pi^2}{2r+1}(r-s)!\,(r+s)!\,(r-t)!\,(r+t)! \quad (3.8)$$

4. We now merely have to follow Schrödinger's method of approximation as applied to nearly degenerate systems. Write

$$\psi_{nku}^{rst}=f_{nk}(r)\,P_k^u(\cos\theta)\,e^{iu\phi}P_r^{s,t}(\cos\chi)\,e^{i(s\lambda+t\mu)}.$$

Then in (2.2)

$$\left(\frac{\partial}{\partial\phi}+2\frac{\partial}{\partial\lambda}\right)\psi_{nku}^{rst}=i(u+2s)\,\psi_{nku}^{rst},$$

and

$$\left[e^{i(\phi-\lambda)}\left\{\frac{\partial}{\partial\theta}+i\cot\theta\frac{\partial}{\partial\phi}\right\}\left\{\frac{\partial}{\partial\chi}+\frac{i}{\sin\chi}\left(\frac{\partial}{\partial\mu}-\cos\chi\frac{\partial}{\partial\lambda}\right)\right\}\right.$$
$$\left.+e^{-i(\phi-\lambda)}\left\{\frac{\partial}{\partial\theta}-i\cot\theta\frac{\partial}{\partial\phi}\right\}\left\{\frac{\partial}{\partial\chi}-\frac{i}{\sin\chi}\left(\frac{\partial}{\partial\mu}-\cos\chi\frac{\partial}{\partial\lambda}\right)\right\}+2\frac{\partial^2}{\partial\phi\,\partial\lambda}\right]\psi_{nku}^{rst}$$
$$=-(k-u)(r+s)\,\psi_{n,k,u+1}^{r,s-1,t}-(k+u)(r-s)\,\psi_{n,k,u-1}^{r,s+1,t}-2us\psi_{nku}^{rst}.$$

Any characteristic function is approximately of the form

$$\psi=\sum_{u,s}a_{us}\psi_{nku}^{rst}.$$

Strictly speaking there should also be a subscript t in a_{us} and a summation

* [*Added in proof.*—This definition applies also when r, s and t are half numbers, and this extends the work to even multiplicities, though of course in these cases the characteristic functions are double-valued.]

for t, but we shall soon see that this is unnecessary. The summations for u and s are restricted to values where both $|u| \leqslant k$ and $|s| \leqslant r$.

Let the associated energy be

$$W = W_{nk} + \frac{h^2}{8\pi^2 I} r(r+1) + \overline{W}.$$

Substituting in (2.2) we get†

$$\overline{W} \sum_{u,s} a_{us} \psi^{rst}_{nku} = \frac{eHh}{4\pi mc} \sum_{u,s} (u+2s)\, a_{us} \psi^{rst}_{nku}$$

$$+ \frac{h^2}{8\pi^2} U(r) \sum_{u,s} a_{us} [(k-u)(r+s)\, \psi^{r,s-1,t}_{u,k,u+1} + (k+u)(r-s)\, \psi^{r,s+1,t}_{n,k,u-1} + 2us \psi^{rst}_{nku}].$$

Multiply by $\psi^{rst}_{nku}{}^*$ and integrate over the whole co-ordinate space. We write $\omega = \dfrac{eHh}{4\pi mc}$ so that ω is the normal Zeeman effect in energy units. Also put

$$\beta = \int \frac{h^2}{8\pi^2} U(r) \{f_{nk}(r)\}^2 r^2 dr \Big/ \int \{f_{nk}(r)\}^2 r^2 dr,$$

β is the constant of the multiplet separation. Then we have

$$\bar{W} a_{us} = \omega (u+2s)\, a_{us} + \beta [(k-u+1)(r+s+1)\, a_{u-1,\, s+1} + (k+u+1)(r-s+1)\, a_{u+1,\, s-1} + 2us a_{us}]. \quad (4.1)$$

This system of difference equations is that which gives the term levels in the Zeeman effect.

The equation (4.1) contains no mention of t, and if t had been added as a subscript to a_{us} we should have obtained identically the same system of equations in a_{ust} for each value of t. These would all give the same levels and same a-ratios, though the characteristic functions would have a very different appearance. We shall see later that this difference is without effect on intensities as well as levels. The apparent complication is only due to the possibility of choosing any arbitrary radius of the sphere as pole for the Eulerian angles.

The fact which simplifies the problem to manageable proportions is that when all equations of this type are set down they fall into "chains" of equations; in each chain the only a's that occur have the sum of their two subscripts a constant, which is in fact the m of the term (while the u and s are the quantities which in a strong field are called m_k and m_r). Examples are given below.

† In this formula r is used in two senses, as the radius in U (r) and elsewhere as a quantum number. It would cause more confusion to alter one of the symbols than to retain this ambiguity.

Each chain gives a determinantal equation of which the roots are the $\overline{W}$'s of the associated levels. We label each of these with an index j (the reason for choice of particular numbers for j only appears later), and then determine the ratios of all the a's for each j. Before exhibiting this and proving its correctness, we shall, however, return to the model and work out intensities.

Schrödinger has shown how this is to be done. Let ψ, ψ' be two characteristic functions, and X any component of electric moment. The associated intensity is

$$I = \left| \int \psi \psi'^* X \right|^2 \Big/ \int |\psi|^2 \int |\psi'|^2, \tag{4.2}$$

where ψ'^* is the complex quantity conjugate to ψ', the integrations being over the whole co-ordinate space. We shall here take the co-ordinates x, y, z of the centre of the body as determining the electric moment. Combining them in the appropriate way, we have to consider the three quantities $x + iy$, $2z$, $x - iy$, or $r \sin\theta\, e^{i\phi}$, $2r \cos\theta$, $r \sin\theta\, e^{-i\phi}$, corresponding respectively to $\perp r \parallel$, $\perp l$, components. If we substitute in (4.2) any of our characteristic functions, the integration for the radius is always the same, so that it falls out of consideration as far as concerns relative intensity, and this is all we shall be concerned with.

Take as our two states those with functions

$$\psi = \sum_s a^{k,j}_{m-s,s}\, \psi^{r,s,t}_{n,k,m-s}$$

and

$$\psi' = \sum_{s'} a^{k',j'}_{m'-s',s}\, \psi^{r,s',t'}_{n',k',m'-s'}.$$

If we multiply ψ by ψ'^* and integrate for μ and λ, the only terms that will not vanish are those for which $t' = t$ and $s' = s$. If then we multiply by, say, $X = r \sin\theta e^{-i\phi}$, the only terms that will not vanish are those for which $m' = m - 1$, and even these will do so unless $k' = k \pm 1$ (see § 3). Similarly if we take $X = 2r \cos\theta$, we must have $m' = m$, $k' = k \pm 1$, and if $X = r \sin\theta e^{i\phi}$, $m' = m + 1$, $k' = k \pm 1$. We thus have three formulæ for intensity, of which the first is

$$k \to k - 1,\ m \to m - 1,\ j \to j'$$

$$\frac{\left| \sum_s a^{k,j}_{m-s,s}\, a^{k-1,j'}_{m-1-s,s}\, \psi^{rst}_{n,k,m-s}\, \psi^{rst*}_{n',k-1,m-1-s} \sin\theta e^{-i\phi} \right|^2}{\sum_s (a^{k,j}_{m-s,s})^2 \left| \psi^{rst}_{n,k,m-s} \right|^2 \sum_s (a^{k-1,j'}_{m-1-s,s})^2 \left| \psi^{rst}_{n',k-1,m-1-s} \right|^2}. \tag{4.3}$$

The integration for μ gives a factor $(r + t)!\,(r - t)!$ by (3.8) in every term of both numerator and denominator. This factor therefore cancels out, and

as the ratios of the a_{us}'s are not affected by t we see that the intensities of the lines, as well as their frequencies, are independent of it. This completes the proof of the statement in § 2 that μ is without effect on the spectrum. Applying to (4.3) the integral formulæ of § 3, and dropping certain constant factors, we have

$$\frac{\left\{\sum_s a^{k,j}_{m-s,s}\, a^{k-1,j'}_{m-1-s,s}\,(r+s)!\,(r-s)!\,(k+m-s)!\,(k-m+s)!\right\}^2}{\left[\left\{\sum_s (a^{k,j}_{m-s,s})^2\,(r+s)!\,(r-s)!\,(k+m-s)!\,(k-m+s)!\right\}\times \left\{\sum_s (a^{k-1,j'}_{m-1-s,s})^2\,(r+s)!\,(r-s)!\,(k+m-s-2)!\,(k-m+s)!\right\}\right]}. \quad (4.4)$$

The other two components are

$$\frac{4\left\{\sum_s a^{k,j}_{m-s,s}\, a^{k-1,j'}_{m-s,s}\,(r+s)!\,(r-s)!\,(k+m-s)!\,(k-m+s)!\right\}^2}{\left[\left\{\sum_s (a^{k,j}_{m-s,s})^2\,(r+s)!\,(r-s)!\,(k+m-s)!\,(k-m+s)!\right\}\times \left\{\sum_s (a^{k-1,j'}_{m-s,s})^2\,(r+s)!\,(r-s)!\,(k+m-s-1)!\,(k-m+s-1)!\right\}\right]} \quad (4.5)$$

for the parallel component, and for the other perpendicular component,

$$\frac{\left\{\sum_s a^{k,j}_{m-s,s}\, a^{k-1,j'}_{m+1-s,s}\,(r+s)!\,(r-s)!\,(k+m-s)!\,(k-m+s)!\right\}^2}{\left[\left\{\sum_s (a^{k,j}_{m-s,s})^2\,(r+s)!\,(r-s)!\,(k+m-s)!\,(k-m+s)!\right\}\times \left\{\sum_s (a^{k-1,j'}_{m+1-s,s})^2\,(r+s)!\,(r-s)!\,(k+m-s)!\,(k-m+s-2)!\right\}\right]}. \quad (4.6)$$

These formulæ give the intensities observed in a direction at right angles to the magnetic field, not the energies emitted.

5. The formulæ (4.1) and (4.4, 5, 6) are all that is required for the complete statement of the standard Zeeman effect. Moreover, though our model fails, the equations derived from it are just as valid for even multiplicities as for odd. For these it is usual to use half quantum numbers, and it is interesting to observe that these half numbers do not occur in the equations at all, but arise in much the same way as does the energy $h\nu\,(n+\frac{1}{2})$ in the problem of a linear vibrator. It is in fact quite easy to formulate the equations without any halves in them at all; it is only necessary to sacrifice symmetry in order to do so, but the statement of conditions becomes somewhat clumsy on account of "end effects," that is, cases where the chain of equations is shorter than it is in general. On this account it is best to make use of half number suffixes for even multiplicities —the coefficients are still all integers. We now restate the rules for finding

the levels and intensities, then give examples, and finally outline the proof that our equations always lead to the observed results.

Take $k = 0, 1, 2 \ldots$ for $s, p, d \ldots$

$r = 0, \frac{1}{2}, 1 \ldots$ for singlets, doublets, triplets . . .

Let u be any integer between $-k$ and k inclusive, and let s be either an integer or half integer with r (*i.e.*, $r + s$ is an integer) and $-r \leqslant s \leqslant r$.

Take any of the permitted values of u and s, and write down the equation

$$-a_{u-1,s+1}\beta(k-u+1)(r+s+1) + a_{u,s}[\overline{W} - \beta \cdot 2us - \omega(u+2s)] - a_{u+1,s-1}\beta(k+u+1)(r-s+1) = 0,$$

in which all the coefficients are integers, whatever r may be. Set down the corresponding equations in $a_{u-1,s+1}$, $a_{u+1,s-1}$, and carry on in both directions until stopped by either of the conditions $|u| \not> k$ or $|s| \not> r$. The determinant of this system of equations will give the values of $\overline{W}$ corresponding to $m = u + s$. If the roots are arranged in decreasing order of magnitude, the greatest will have $j = k + r$, and the others are to be numbered by units downwards as far as the chain goes. u and s are themselves m_k and m_r in strong fields. The intensity of a line is given by applying the formulæ of § 4 (and we may again note that all the coefficients are integers there).

As it is easier to follow arithmetic than algebra, we shall give an example, and shall take $k < r$. To avoid having to write an enormous number of equations we shall take the p-levels of the quartet system.

Quartet system, p-terms. $k = 1, r = 3/2$.

The chains are

$$(1) \quad a_{1,3/2}[\overline{W} - 3\beta - 4\omega] = 0,$$

$$(2) \quad a_{0,3/2}[\overline{W} - 3, \omega] - a_{1,1/2}\beta \cdot 2 \cdot 1 = 0,$$
$$-a_{0,3/2}\beta \cdot 1 \cdot 3 + a_{1,1/2}[\overline{W} - \beta - 2\omega] = 0,$$

$$(3) \quad a_{-1,3/2}[\overline{W} + 3\beta - 2\omega] - a_{0,1/2}\beta \cdot 1 \cdot 1 = 0,$$
$$-a_{-1,3/2}\beta \cdot 2 \cdot 3 + a_{0,1/2}[\overline{W} - \omega] - a_{1,-1/2}\beta \cdot 2 \cdot 2 = 0,$$
$$-a_{0,1/2}\beta \cdot 1 \cdot 2 + a_{1,-1/2}[\overline{W} + \beta] = 0,$$

$$(4) \quad a_{-1,1/2}[\overline{W} + \beta] - a_{0,-1/2}\beta \cdot 1 \cdot 2 = 0,$$
$$-a_{-1,1/2}\beta \cdot 2 \cdot 2 + a_{0,-1/2}[\overline{W} + \omega] - a_{1,-3/2}\beta \cdot 2 \cdot 3 = 0,$$
$$-a_{0,-1/2}\beta \cdot 1 \cdot 1 + a_{1,-3/2}[\overline{W} + 3\beta + 2\omega] = 0,$$

$$(5) \quad a_{-1,-1/2}[\overline{W} - \beta + 2\omega] - a_{0,-3/2}\beta \cdot 1 \cdot 3 = 0,$$
$$-a_{-1,-1/2}\beta \cdot 2 \cdot 1 + a_{0,-3/2}[\overline{W} + 3\omega] = 0,$$

$$(6) \quad a_{-1,-3/2}[\overline{W} - 3\beta + 4\omega] = 0.$$

If we form the determinant for each chain, we get an algebraic equation which determines one, two, or three of the levels, for any value of $\omega : \beta$. For the general case numerical solutions would have to be used; but if ω is much greater than β, we can immediately see that all the levels are approximately multiples of ω—the Paschen-Back effect. Though our formulæ are valid in all cases, general algebraic formulæ are not easy to follow, and we therefore discuss the extreme cases, which were the only ones treated by earlier writers. To order the levels for weak fields we approximate with ω small. For example, the chain (2) gives, for $m = 3/2$

$$(\overline{W} - 3\omega)(\overline{W} - \beta - 2\omega) = 6\beta^2$$

whence

$$\overline{W} = 3\beta + \frac{12}{5}\omega, \quad -2\beta + \frac{13}{5}\omega.$$

These solutions we label respectively $j = 5/2$, $j = 3/2$, according to our rule. Furthermore for no value of ω are the roots equal, so that as ω increases adiabatically the roots cannot cross, and we can associate together the solutions in weak and strong fields. The first here becomes 3ω, the second $2\omega + \beta$. The following table shows the whole solution :—

Chain.	$\overline{W}$ weak.	$\overline{W}$ strong.	m.	j.	Weak fields, "a" ratios.
1	$3\beta + 4\omega$	$4\omega + 3\beta$	5/2	5/2	1
2	$3\beta + \frac{12}{5}\omega$	3ω	3/2	5/2	2 3
	$-2\beta + \frac{13}{5}\omega$	$2\omega + \beta$	3/2	3/2	1 −1
3	$3\beta + \frac{4}{5}\omega$	$2\omega - 3\beta$	1/2	5/2	1 6 3
	$-2\beta + \frac{13}{15}\omega$	ω	1/2	3/2	1 1 −2
	$-5\beta + \frac{4}{3}\omega$	$-\beta$	1/2	1/2	1 −2 1
4	$3\beta - \frac{4}{5}\omega$	$-\beta$	−1/2	5/2	3 6 1
	$-2\beta - \frac{13}{15}\omega$	$-\omega$	−1/2	3/2	−2 1 1
	$-5\beta - \frac{4}{3}\omega$	$-2\omega - 3\beta$	−1/2	1/2	1 −2 1

Chain.	$\overline{W}$ weak.	$\overline{W}$ strong.	m.	j.	Weak fields, "a" ratios.	
5	$3\beta - \frac{12}{5}\omega$	$-2\omega + \beta$	$-3/2$	$5/2$	3	2
	$-2\beta - \frac{13}{5}\omega$	-3ω	$-3/2$	$3/2$	-1	1
6	$3\beta - 4\omega$	$-4\omega + 3\beta$	$-5/2$	$5/2$	1	

It may be well here to mention a few points about the chains which are not entirely obvious from the algebra, but which were in fact found most useful in constructing them. We can show that many of the coefficients are immediately determined simply by the existence of the chain type of equations. Take first the case of $k > r$. Then the first member of a chain will involve terms in a_{ur} $(\overline{W} - \ldots)$ and $a_{u+1,\,r-1}$, and the second in a_{ur}, $a_{u+1,\,r-1}$ $(\overline{W} - \ldots)$, $a_{u+2,\,r-2}$. Now if the coefficients are to be algebraic, they must be such that if we put $u = k$ we get a chain of one member, for the first equation is then reduced to the one term in a_{kr} $(\overline{W} - \ldots)$. It follows that in the second equation the coefficient of a_{ur} must contain a factor $(k-u)$, in order that this equation may be satisfied identically for $u = k$. Similar considerations show that in the third equation the coefficient of $a_{u+1,\,r-1}$ must have a factor $k-u-1$, in order that it may vanish identically for the case $u = k - 1$, which is required so as to give the two-member chain. The argument continues all down the chain and determines a factor for each member lying on one side of the diagonal. By attacking the chain from the lower end, and putting $u = -k$, etc., similar factors can be deduced for the other side of the diagonal. The other factors in $r + s + 1$, etc., follow out of the consideration of cases where $k < r$, so that k and not r limits the length of the chain. This argument is of course not conclusive, but the process indicated was in fact found to be most useful in constructing the equations.

Our example of the quartet p-terms has shown how the chains go for $k < r$. When $k > r$ the corresponding arrangement of chains is rather similar. It starts with chains of 1, 2, 3 . . . members, increasing up to the maximum multiplicity, stays there, and then decreases back to 1. The j's of the end chains go from $k + r$ downwards; in the middle part they go down to $k - r$. Thus for example the chains of quartet d are of lengths 1, 2, 3, 4, 4, 3, 2, 1 members, while for quintet f they would be of 1, 2, 3, 4, 5, 5, 5, 4, 3, 2, 1 members.

With a view to illustrating the matter further, and also to calculating some intensities, we will give one of the four-chains of quartet d. Take that for which $m = \frac{1}{2}$ (and of course $k = 2$, $r = 3/2$)

$$\begin{aligned}
&a_{-1,\frac{3}{2}}[\overline{W} + 3\beta - 2\omega] - a_{0,\frac{1}{2}}\beta \,.\, 2 \,.\, 1 && = 0,\\
-a_{-1,\frac{3}{2}}\beta \,.\, 3 \,.\, 3 \;+a_{0,\frac{1}{2}}[\overline{W} - \omega] \quad &- a_{1,-\frac{1}{2}}\beta \,.\, 3 \,.\, 2 && = 0.\\
-a_{0,\frac{1}{2}}\beta \,.\, 2 \,.\, 2 \quad &+ a_{1,-\frac{1}{2}}[\overline{W} + \beta] - a_{2,-\frac{3}{2}}\beta \,.\, 4 \,.\, 3 && = 0.\\
&- a_{1,-\frac{1}{2}}\beta \,.\, 1 \,.\, 1 + a_{2,-\frac{3}{2}}[\overline{W} + 6\beta + \omega] && = 0.
\end{aligned}$$

For weak fields the roots are

$$\overline{W} = 6\beta + \tfrac{5}{7}\omega, \quad -\beta + \tfrac{24}{35}\omega, \quad -6\beta + \tfrac{3}{5}\omega, \quad -9\beta.$$

So as to give an example of the calculation of intensities, we may note that for the first solution ($j = 7/2$)

$$a_{-1,\frac{3}{2}} = 4, \quad a_{0,\frac{1}{2}} = 18, \quad a_{1,-\frac{1}{2}} = 12, \quad a_{2,-\frac{3}{2}} = 1.$$

In applying the formulæ for intensities, we have to remember that they are all relative, so that we shall have to work out two. Now of the whole multiplet *line*—and this is true of all cases—two members have specially simple behaviour, and these are the lines connecting respectively the two first and the two last chains, each of which has only one member. These lines behave as though the Zeeman effect had no anomaly; they keep their intensity constant, and their shift increases uniformly with the magnetic field according to the strict Larmor theorem. We shall call them the leading members of the multiplet. For quartet d the leading chain gives

$$k = 2, \; r = \tfrac{3}{2}, \; m = \tfrac{7}{2}, \; j = \tfrac{7}{2}, \; \overline{W} = 6\beta + 5\omega, \; a_{2,\frac{3}{2}} = 1,$$

so the intensity of the leading member of the pd lines is

$$\{3\,!\;0\,!\;4\,!\;0\,!\}^2 / \{3\,!\;0\,!\;4\,!\;0\,!\}\{3\,!\;0\,!\;2\,!\;0\,!\} = 12.$$

Compare with this the intensity in a weak field of the line

$$k, 2 \to 1; \quad j, \tfrac{7}{2} \to \tfrac{5}{2}; \quad m, \tfrac{1}{2} \to -\tfrac{1}{2}.$$

Our formula gives

$$\begin{aligned}
&\{0 + 18\,.\,3\,.\,2\,!\;1\,!\;2\,!\;2\,! + 12\,.\,6\,.\,1\,!\;2\,!\;3\,!\;1\,! + 1\,.\,1\,.\,0\,!\;3\,!\;4\,!\;0\,!\}^2\\
&\div[\{4^2\,.\,3\,!\;0\,!\;1\,!\;3\,! + 18^2\,.\,2\,!\;1\,!\;2\,!\;2\,! + 12^2\,1\,!\;2\,!\;3\,!\;1\,! + 1^2\,.\,0\,!\;3\,!\;4\,!\;0\,!\}\\
&\times\{3^2\,.\,2\,!\;1\,!\;0\,!\;2\,! + 6^2\,.\,1\,!\;2\,!\;1\,!\;1\,! + 1^2\,.\,0\,!\;3\,!\;2\,!\;0\,!\}] = \tfrac{24}{7},
\end{aligned}$$

i.e., 2/7 of the leading line.

We will also verify an example of the fact that in weak fields lines involving $j \to j - 2$ give no intensity, a result by no means obvious with the present method. Take the line

$$k, 2 \to 1; \; j, \tfrac{7}{2} \to \tfrac{3}{2}; \; m, \tfrac{1}{2} \to -\tfrac{1}{2}.$$

We only need the numerator. It is

$$\{18\,(-2)\,.\,2!\;1!\;2!\;2!+12\,.\,1\,.\,1!\;2!\;3!\;1!+1\,.\,1\,.\,0!\;3!\;4!\;0!\}^2$$
$$=\{-288+144+144\}^2 \qquad =0.$$

6. There are a good many end cases that would have to be treated to construct the complete proof that we have really got the standard multiplet. We shall, however, be content to treat here only of the " complete " chains, which are those where $k \geqslant r$, while $m \leqslant k-r$. The reader can deal with the other cases for himself by similar methods.

Since the sum of the suffixes in a chain is always the same, we shall drop one of them, retaining s, which for the case we consider runs from r to $-r$. We shall set down the chain for m. It is

$$a_r\,[\overline{\mathrm{W}}-\beta\,.\,2r\,(m-r)-\omega\,(m+r)]-a_{r-1}\,\beta\,(k+m-r+1)\,.\,1 \qquad =0$$

$$-a_r\,\beta\,(k-m+r)\,2r+a_{r-1}\,[\overline{\mathrm{W}}-\beta\,(2r-2)\,(m-r+1)-\omega\,(m+r-1)]$$
$$-a_{r-2}\,\beta\,(k+m-r+2)\,.\,2 \qquad =0$$

. . . .

and are

$$b_r\{\overline{W} - \beta \cdot 2kr - \omega(m+r)\} + b_{r-1}\omega = 0$$

$$-b_r\beta(k-m+r)\,2r + b_{r-1}\{\overline{W} - \beta \cdot 2(kr-k-r) - \omega(m+r-1)\} + b_{r-2} \cdot 2\omega = 0$$

$$\cdots\cdots \qquad \cdots\cdots \qquad \cdots\cdots$$

$$-b_{j-k+1}\beta(j-m+1)(j-k+r+1) + b_{j-k}\{\overline{W} - \beta[j(j+1) - k(k+1) - r(r+1)] - \omega(m+j-k)\} + b_{j-k-1}(k+r-j+1)\omega = 0$$

$$\cdots\cdots \qquad \cdots\cdots \qquad \cdots\cdots$$

From these equations it is evident that if $\omega = 0$, we have Landé's γ formula because the determinant will have no members on one side of the diagonal. If we put $\beta = 0$, we have the Paschen-Back effect as all the members on the other side of the diagonal now vanish, but of course our transformation was unnecessary in order to show this.

It remains to prove Landé's g-formula, and to do this we solve first for the b's with $\omega = 0$ and a given j. It then appears that all the b's with index greater than $j - k$ vanish, while the rest are easily expressed as products of two binomial coefficients. When ω is not quite zero, the upper b's no longer quite vanish, but it is easy to see that each bears to the next a ratio of order ω. Let the characteristic required be

$$\overline{W} = \beta[j(j+1) - k(k+1) - r(r+1)] + q\omega.$$

Then the value of q will be obtained from the three equations

$$b_{j-k+1}\{\beta j(j+1) - \beta(j+1)(j+2)\} + b_{j-k}(k+r-j)\,\omega = 0$$

$$-b_{j-k+1}\beta(j-m+1)(j-k+r+1) + b_{j-k}(q-m-j+k)\,\omega + b_{j-k-1}(k+r-j+1)\,\omega = 0.$$

$$-b_{j-k}\beta(j-m)(j-k+r) + b_{j-k-1}\{\beta j(j+1) - \beta(j-1)j\} = 0.$$

The determinant of these equations reduces to Landé's g-formula,

$$q = m\left\{1 + \frac{j(j+1) - k(k+1) + r(r+1)}{2j(j+1)}\right\}$$

If the algebra of the transformation is examined, it appears that at no stage of it is it necessary to assume how many members the chain of equations has. Consequently the proof is valid for the incomplete end chains as well as for the middle ones. It would only require trivial modifications to make it applicable for $k < r$.

The connection of the roots with those in the Paschen-Back effect is obvious, for we can order them together from the top of the chain in each case. This

rests on the fact that for no value of ω can the determinant of a chain have equal roots, as may easily be shown by a Sturmian method, so that as ω increases adiabatically, the root which starts as, say, p^{th} in order must end as p^{th} too. For example, the first solution, which for small fields is $\beta \,.\, 2kr + m\left(1+\frac{r}{k+r}\right)\omega$, becomes in strong fields $(m+r)\,\omega + \beta \,.\, 2r\,(m-r)$. In tracing out the connections in general, it must be remembered that there is an essential asymmetry between the top and bottom of the chain system.

It is easy to solve for the ratios of the b's for zero field with a given j, and thence to deduce the corresponding a's; but the expressions are not very simple. They are required in order to determine intensities in weak fields. I have, as a matter of fact, examined the general case sufficiently to see that it is only a matter of heavy algebraic manipulation, without inherent difficulty, to simplify these intensity formulæ; but it seemed sufficient to be content with actually calculating them only for a number of cases (*e.g.*, quintet system, arbitrary k and m). Wherever this has been done, the values given by Kronig* are verified.

In this connection, however, there is one very important consideration that cannot be neglected—the combination rule for j. This also is of course verified where it has been tried, but the type of calculation required to do so is so intricate that it is not in the least suggestive of a general principle, but rather of a sort of accident; so that even if we gave a proof starting from the formulæ of § 4 it would not be very convincing. The plain fact is that we have been treating of all field strengths indifferently, and except in the case of zero field j is nothing more than a name. To fit in with the familiar notation, we have adopted numeration by j, instead of numbering from say 0 to $2r$, and the consequence is merely a slight simplification of the algebra. To bring out the force of j we shall outline a new attack on the whole problem, omitting entirely the terms introduced on account of the magnetic field.

7. The quantum number j only has a real meaning in the absence of the external magnetic field, so that we may say that it depends on considerations involving the isotropy of the atom. Now this isotropy could hardly be better concealed than by spherical harmonics in the form in which we have used them. To exhibit it we must adopt a different method of expressing them, and this is not hard, at any rate in the case of odd multiplicities. We saw that the characteristic functions were of the form

$$f_{nk}(r)\,\mathrm{P}_k^m(\cos\theta)\,e^{im\phi}\,\mathrm{P}_r^{s,\,t}(\cos\chi)\,e^{i(s\lambda+t\mu)},$$

* 'Z. f. Physik,' vol. 31, p. 885 (1925).

and that all the results were unaffected by t. Put $t = 0$, and we have simply another ordinary spherical harmonic $P_r^s(\cos\chi)\, e^{is\lambda}$.

Now any spherical harmonic in x, y, z can be written as a combination of terms

$$r^{a+b+c+1}\left(\frac{\partial}{\partial x}\right)^a\left(\frac{\partial}{\partial y}\right)^b\left(\frac{\partial}{\partial z}\right)^c\frac{1}{r}, \tag{7.1}$$

and in this form the isotropy can be made explicit. We also introduce three co-ordinates ξ, η, ζ, with radius ρ and replace $P_r^s(\cos\chi)\, e^{is\lambda}$ by

$$\rho^{a'+b'+c'+1}\left(\frac{\partial}{\partial \xi}\right)^{a'}\left(\frac{\partial}{\partial \eta}\right)^{b'}\left(\frac{\partial}{\partial \zeta}\right)^{c'}\frac{1}{\rho}. \tag{7.2}$$

The operator occurring in the last line of (2.2) is then

$$-\frac{h^2}{4\pi^2}U(r)\left\{\left(y\frac{\partial}{\partial z}-z\frac{\partial}{\partial y}\right)\left(\eta\frac{\partial}{\partial \zeta}-\zeta\frac{\partial}{\partial \eta}\right)+\left(z\frac{\partial}{\partial x}-x\frac{\partial}{\partial z}\right)\left(\zeta\frac{\partial}{\partial \xi}-\xi\frac{\partial}{\partial \zeta}\right)+\left(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}\right)\left(\xi\frac{\partial}{\partial \eta}-\eta\frac{\partial}{\partial \xi}\right)\right\},$$

or say $\left(\left[r\cdot\frac{\partial}{\partial r}\right]\left[\rho\cdot\frac{\partial}{\partial \rho}\right]\right)$ in vector notation. This operator commutes with r and ρ, and as before we can omit consideration of the radius while ρ is purely auxiliary, and plays no part in the integrations. We can drop out the factors $r^{a+b+c+1}$ and $\rho^{a'+b'+c'+1}$ from (7.1) and (7.2) as of no interest, and can take as the equation to determine the characteristic values,

$$\overline{W}\psi = -2\beta\left(\left[r\cdot\frac{\partial}{\partial r}\right]\left[\rho\cdot\frac{\partial}{\partial \rho}\right]\right)\psi,$$

where in order to satisfy the nul approximation ψ must be a sum of products of harmonics in x, y, z of degree k and in ξ, η, ζ of degree r. Now the operator in this equation is invariant for a change of axes, provided that the transformation is applied simultaneously to x, y, z, and ξ, η, ζ. Hence any solution that can be obtained must be of tensor form. We shall show that *j is the rank of the tensor solution.*

The present section does not pretend to completeness, and indeed we shall only give a few examples to bring out this interpretation of j. It will be convenient to use a tensor notation, so we shall write x_a for x, y, or z, and ξ_a for ξ, η, ζ. We follow the tensor notation in omitting the summation sign when duplicated indices are to be summed; but there is here no distinction between co- and contravariance, so that all the indices are written below. We shall make use of the notation of vector products with suffix for the component.

Thus $[x\xi]_1 = x_2\xi_3 - x_3\xi_2$. Any spherical harmonic, *e.g.*, $\frac{\partial}{\partial x_\alpha}\frac{\partial}{\partial x_\beta}\frac{\partial}{\partial x_\gamma}\cdot\frac{1}{r}$ we write simply as $d_\alpha\, d_\beta\, d_\gamma$. Similarly for $\frac{\partial}{\partial \xi_\alpha}\frac{\partial}{\partial \xi_\beta}\frac{\partial}{\partial \xi_\gamma}\cdot\frac{1}{\rho}$ we write $\mathrm{D}_\alpha\, \mathrm{D}_\beta\, \mathrm{D}_\gamma$.

The equation for the characteristic functions is then

$$\begin{aligned}\overline{\mathrm{W}}\psi &= -2\beta\left[x\frac{\partial}{\partial x}\right]_a\left[\xi\frac{\partial}{\partial \xi}\right]_a\psi \\ &= -\beta\left\{\left(x_\beta\frac{\partial}{\partial x_\gamma} - x_\gamma\frac{\partial}{\partial x_\beta}\right)\left(\xi_\beta\frac{\partial}{\partial \xi_\gamma} - \xi_\gamma\frac{\partial}{\partial \xi_\beta}\right)\right\}\psi.\end{aligned}$$

The appropriate solution will be a sum of products each of which has k d's and r D's.

Consider the triplet p terms. The permissible forms of solution are $\psi = d_\alpha \mathrm{D}_\beta$, and we have to find what combination gives an actual solution. Solving out in detail (either directly or by transformation of the work of § 4), we find the following solutions

$$\begin{aligned}\overline{\mathrm{W}} &= -4\beta \quad j = 0 \quad \psi = \mathrm{C} d_\alpha \mathrm{D}_\alpha \\ \overline{\mathrm{W}} &= -2\beta \quad j = 1 \quad \psi = \mathrm{C}_\alpha [d\mathrm{D}]_\alpha \\ \overline{\mathrm{W}} &= \quad 2\beta \quad j = 2 \quad \psi = \mathrm{C}_{\alpha\beta} d_\alpha \mathrm{D}_\beta\end{aligned}$$

$$\text{with} \quad \mathrm{C}_{\alpha\beta} = \mathrm{C}_{\beta\alpha}, \quad \text{and} \quad \mathrm{C}_{\alpha\alpha} = 0.$$

The C's here are arbitrary constants, and the conditions attached to the last, which reduce them from 9 to 5, are just those that exclude terms corresponding to the previous solutions. If we do the same work for the triplet d terms, the typical solution must depend on $d_\alpha\, d_\beta\, \mathrm{D}_\gamma$, and it is possible to build three tensors out of this, one of the third rank, one of the second, by what we may call curl-contraction, and one of the first by complete contraction. The solutions are

$$\begin{aligned}\overline{\mathrm{W}} &= -6\beta \quad j = 1 \quad \psi = \mathrm{C}_\alpha d_\alpha d_\beta \mathrm{D}_\beta \\ \overline{\mathrm{W}} &= -2\beta \quad j = 2 \quad \psi = \mathrm{C}_{\alpha\beta} d_\alpha [d\mathrm{D}]_\beta\end{aligned}$$

$$\text{with} \quad \mathrm{C}_{\alpha\beta} = \mathrm{C}_{\beta\alpha}, \quad \mathrm{C}_{\alpha\alpha} = 0.$$

$$\overline{\mathrm{W}} = 4\beta \quad j = 3 \quad \psi = \mathrm{C}_{\alpha\beta\gamma} d_\alpha d_\beta \mathrm{D}_\gamma$$

$$\text{with} \quad \mathrm{C}_{\alpha\beta\gamma} = \mathrm{C}_{\beta\alpha\gamma} = \mathrm{C}_{\alpha\gamma\beta} \quad \text{and} \quad \mathrm{C}_{\alpha\beta\beta} = 0.$$

It is easy to enumerate the numbers of independent solutions that these give. They are of course 3, 5, 7.

Such examples can be multiplied indefinitely, but perhaps the above will suffice. For a complete discussion it would be necessary to elaborate somewhat

the notation and method of this section, and this would open a wider field than the present paper contemplates. In any particular case it is fairly easy to verify the combination law of j, exhibiting it as dependent on the tensor-ranks of the two characteristic functions. We shall not go further into the matter here.

In conclusion we may note a few points of general interest. We have succeeded in giving to j a very clear physical meaning, but only for odd multiplets. There seems no simple way of extending the idea to even—to do so would require the invention of tensors of half rank! So here again we come to the difficulty of the half quantum number of the spinning electron.

The model that we have used gives, of course, only a very abbreviated account of the interactions of a group of electrons, and it is doubtful how far its extension is legitimate. For instance, consider the case of a triplet-singlet intercombination line. To give such a line with our model it is only necessary to extend the idea of electric moment. It is easy to show that every feature of the intercombination* is given by replacing the vector x_a in the electric moment by $[x\xi]_a$. We may observe that the presence of ξ in the quasi-moment seems natural, for Heisenberg has explained that singlet and triplet in helium would not combine at all but for the magnetism in the electron, and so we should expect the intensity of an intercombination line to involve the spin.

Summary.

The problem of a spinning electrified body moving in a central orbit in a magnetic field is solved by the method of the wave mechanics in spherical harmonics. It is shown to lead to a set of simple arithmetical equations which exactly give all the features of the standard Zeeman effect in all strengths of field. The model only yields strictly the odd multiplicities, but the same system of equations is just as competent to give the even. Formulæ are given for the intensity of any line at any strength of field. A few examples are worked out. The development in spherical harmonics brings out strongly the meaning of the quantum numbers k, m, r, m_k, m_r, but much more obscurely, j; and an alternative development is sketched which brings out the full force of j.

I must express my thanks to Dr. R. Schlapp for his help in working out several of the details.

* Except the *absolute* values of the intensities.

Electron Emission under the Influence of Chemical Action at Higher Gas Pressures, and some Photo-electric Experiments with Liquid Alloys.

By O. W. Richardson, F.R.S., Yarrow Research Professor of the Royal Society, and M. Brotherton, Ph.D., King's College, London.

(Received March 31, 1927.)

§ 1. *Apparatus and Procedure.*

The present paper deals with the emission of electrons from metallic drops in the presence of a gaseous chemical reagent and is a continuation of work described in four earlier publications by one or other of us.* These will be referred to throughout this paper by the numbers assigned to them in the present reference. In (3) the phenomena were studied at what are believed to be very low partial pressures of the reactive gas which was controlled by manipulating a bath of liquid air round the tube containing this gas in a condensed phase. This control is very inexact and many of the measurements were made under conditions in which the density of the gas was believed to be varying rapidly with time. This necessitates a large number of check experiments and is conducive to errors. In (4) methods were devised for maintaining the pressure of the reactive gas at a sufficiently constant value over a period required to take a succession of observations. These methods have since been further developed and it is believed that the pressure is now under sufficiently reliable control down to partial pressures as low as 0·001 mm. The present paper deals with the phenomena as they manifest themselves with pressures of reactive gas which are not less than about 0·001 mm. The reaction studied throughout has been that of $COCl_2$ on liquid alloys of Na and K. The composition of the alloys has been varied very considerably.

The testing chamber is the same as that described in (3) (p. 23) except that (*a*) the reservoir containing the supply of alloy is connected to the testing chamber by a ground glass joint. This facilitated the dismantling and putting together of the testing chamber. (*b*) The copper rod supporting the electrode is surrounded by a quartz tube and then by a layer of copper wire which is

* (1) O. W. Richardson, 'The Emission of Electricity from Hot Bodies,' 1st Edn., p. 293, London (1916). (2) O. W. Richardson, 'The Emission of Electricity from Hot Bodies,' 2nd Edn., p. 310, London (1921). (3) O. W. Richardson, 'Phil. Trans., A, vol. 222, p. 1 (1921). (4) M. Brotherton, 'Roy. Soc. Proc.,' A, vol. 105, p. 468 (1924).

earthed. This is to prevent stray charges from leaking on to the electrode. (*c*) The lower reservoir is closed by a brass cap sealed to the apparatus with sealing wax. This greatly facilitated the removal of the alloy when necessary and enabled it to be connected to earth by a wire passing through the brass cap.

The arrangement for supplying the active gas was the same as that described in (4), p. 473, viz. :—by diffusion from a reservoir into the exhausted testing apparatus through a fine nozzle. The pressure in the stream of gas may be varied either by varying the density of the gas in the reservoir, or by varying the size of the nozzle, but it is obvious that the most constant pressures will be obtained by using large masses of gas and fine nozzles. In most of the experiments the exhaustion was by a Gaede mercury pump run continuously at a constant rate. In some this was replaced by a tube containing charcoal immersed in liquid air. This is a very good method for getting the lowest pressures in the range here dealt with. If, during an experiment, the pressure begins to fall, the matter can be put right by passing a little gas into the reservoir. The quantity of gas in this is easily adjustable, any excess being readily re-condensed into the generator by means of liquid air.

As a pressure indicator the MacLeod gauge was abandoned in favour of a Pirani gauge supplied by the General Electric Co., Ltd., which was inserted as compactly as possible between the nozzle and the testing chamber. It is not certain that this will give the pressure at the surface of the reacting alloy correctly owing to the presence of the gaseous products of reaction (presumably CO), but it should give a correct indication of the partial pressure of the $COCl_2$ as it streams into the vessel containing the alloy. However, if the reaction is of the straightforward type, as for example that represented by the equation $Na\,K + COCl_2 = Na\,Cl + K\,Cl + CO$, the reaction will not alter the gas pressure, 1 molecule of CO being liberated for every molecule of $COCl_2$ taken up. In any event the Pirani gauge has the great advantage that it indicates the pressure of the gas at any instant, and whether it is steady, or if it is varying, the rate and magnitude of the variations. It is very important to have a knowledge of these factors. The Pirani gauge was tested against a MacLeod gauge and found to agree with the formula supplied with it, using suitable constants, at any rate for pressures below 0·07 mm.

One of the great difficulties in this investigation, as has been realised from the first (see for example (3), p. 9), arises from a kind of hysteresis which sometimes appears in the phenomena. This will be dealt with more fully in § 10 below, but it is necessary to mention it here. For most of the experiments it is essential to get rid of this effect or at least to reduce its consequences to negligible

proportions. Sometimes it has been possible to eliminate this effect by taking down the apparatus and cleaning the outer electrode, but in other cases this treatment made little or no difference. It has been found possible to obtain consistent and reproducible results by applying a small saturating potential, generally 4 volts, between each current reading under a lower potential. This has the effect of making the voltages for a given current below saturation (accelerating voltages being considered positive) to be higher than would otherwise be found, and the result may be to increase the contact potential difference as measured from the chemical emission (see § 4 below) by 0·1 volt in extreme cases. It is not believed that these small displacements affect the slope (see § 4) to a measurable extent. In fact in some of the experiments the displacements were entirely absent and the slope was found to be unaffected.

§ 2. *Effect of the Rate of Drops on the Emission.*

With an alloy of given composition and with the gas at a fixed pressure the electron emission increases rapidly as the number of drops per second increases. The values of the saturation current in Table I were measured for different rates of drops for the alloy Na K at a constant pressure of $COCl_2$ of 0·0065 mm. These drops were all of about the same size.

Table I.

Time of 1 drop (secs.)	7·5	9·50	13·4	114·0
Saturation current (amps.)	$1·0 \times 10^{-8}$	$5·1 \times 10^{-9}$	$3·33 \times 10^{-9}$	$2·33 \times 10^{-10}$

These results can be considered fairly typical of the whole of the observations. We have not measured the size of the drops but all our observations are in qualitative agreement with the view that the number of electrons emitted by an alloy of given composition is proportional to the area of the surface generated in the gas.

§ 3. *Effect of Pressure.*

When the rate and size of the drops is kept constant the electron emission is independent of the pressure of the $COCl_2$ over a range of pressure from 0·005 to 0·05 mm. This is shown by Tables II and III which represent two sets of experiments with different rates of drops, the rate and size of the drops being as constant as possible in each set of experiments.

Table II.

Pressure of $COCl_2$ (mms. of Hg)	0·0065	0·014	0·019	0·028	0·041
Saturation current(amps.)	$5·1 \times 10^{-9}$	$4·75 \times 10^{-9}$	$4·75 \times 10^{-9}$	$4·83 \times 10^{-9}$	$4·50 \times 10^{-9}$
Time of 1 drop (secs.)	9·50	10·1	9·66	10·0	10·5

Table III.

Pressure of $COCl_2$ (mms. of Hg)....	0·005	0·0075	0·011	0·014	0·020	0·024	0·029
Saturation current (amps. $\times 10^{-9}$)	5·41	4·75	4·66	4·50	4·54	4·66	4·58
Time of 1 drop (secs.)	6·44	6·44	6·50	6·50	6·50	6·66	6·72

These experiments were all made with an alloy of the composition Na K. It will be noticed that the saturation currents are not quite constant but seem to diminish slightly as the pressure rises. Any part of this change which is real can be accounted for by the change in the time of the drops which tend to become slower at the higher pressures. The constancy of the emission over this wide range of gas pressure is surprising in view of the results described in (3) where the control of the emission by immersion of the source of $COCl_2$ in a liquid air bath, and the rapid changes of the emission with time, are only comprehensible if the emission was a function of the partial pressure of the $COCl_2$. However, as the emission is constant for pressures between 0·005 and 0·050 mm. and is zero for zero pressure, there must be some intervening range of pressure within which it is a function of the pressure. The present observations suggest that this range lies below 0·001 mm. which is about the lower limit of the pressures measurable on the Pirani gauge employed in these experiments.

§ 4. *The Energy of the Emitted Electrons.*

In (3) it was found that to the degree of accuracy of the tests which could be imposed the electrons were emitted from the drops of Na K_2 with a Maxwell distribution of energy for a temperature of about 3300° K. If the electrons have this distribution of energy, in the case of a small source, such as a drop, at the centre of a large spherical electrode, the current i against an opposing potential V is given by

$$i = i_0 (1 + \sigma V) e^{-\sigma V}, \qquad (1)$$

where i_0 is the saturation current and

$$\sigma = e/kT. \qquad (2)$$

In equation (1) V is the potential difference between points just outside the surfaces of the two electrodes and is not directly accessible to measurement, the experimental scale of potential V_1 being that given by the readings of a voltmeter with its terminals in metallic connection with the two electrodes. As a matter of fact

$$V = V_1 + K \tag{3}$$

where K is the contact potential difference between the two surfaces. We avoid this difficulty by making use of the fact, shown by equation (1), that i/i_0 is a function of σV only. By using the data in Table IX, p. 36 of (3) we construct a large scale graph of $\log i/i_0 (= (1 + \sigma V) e^{-\sigma V})$ against $\sigma V (= eV/kT)$. This graph which is almost a straight line except near the origin represents a relation between $\log i/i_0$ and σV which is true generally apart from any particular value of σ or T. With the help of this graph we read off the values of σV which correspond to the experimental values of $\log i/i_0$. The values of V_1 to which $\log i/i_0$ correspond are known from the experiments, and we now construct a new graph in which the values of σV are plotted against the corresponding values of V_1. If equations (1) and (3) are true this should be a straight line which intersects the V_1 axis at $V = 0$, *i.e.*, at $V_1 = -K$ and whose "slope" σ is equal to e/kT. Thus the equivalent temperature of the emitted electrons is

$$T = \frac{e}{k\sigma} = \frac{1{\cdot}1887 \times 10^4}{\text{slope}} {}^\circ\text{K}, \tag{4}$$

if V_1 is in volts and the values taken from (3) Table IX, where eV/kT is in E.S. units are used. The contact P.D. is the intercept on the V_1 axis between the slope line and the origin. It is not necessary to determine this intersection directly. Having first determined σ it can be found from the second graph by noting the value of σV corresponding to any particular V_1 dividing it by σ and adding the two voltages together.

It is perhaps as well to state at the outset that the results obtained are in harmony with equations (1)–(3) except in the neighbourhood of $V = 0$. The slopes obtained show very little variation over a very wide range of variation in the other factors. They are, however, uniformly higher than the slopes which were found in (3) and correspond to a temperature in the neighbourhood of 2300° K. or about 1000° less than that given by the former experiments.

Out of a very large number of experiments of this type we shall select one as typical for detailed consideration. The one selected was made with an alloy of composition Na K. The drops were slow (1 drop in 17·7 secs.). The exhaustion was by liquid air and charcoal and the partial pressure of the $COCl_2$

less than 0·001 mm. of mercury. The saturation current was $2{\cdot}58 \times 10^{-9}$ amp. The values i/i_0 at different accelerating voltages (V_1) are given in Table IV.

Table IV.

Volts (V_1)	0·2	0·4	0·6	0·8	1·0	1·2	1·4	1·6	1·8
i/i_0	0·000197	0·000658	0·00150	0·00368	0·00989	0·0216	0·0495	0·1078	0·1605

continued

Volts (V_1)	2·0	2·2	2·4	2·6	2·8	3·0	3·2	3·4	3·8
i/i_0	0·2631	0·366	0·495	0·653	0·763	0·855	0·882	0·947	1·00

These data are plotted in fig. 1. This curve is hardly distinguishable by visual examination from any of the curves between i/i_0 and V_1 described in § 4–§ 9 of this paper, except for some displacements along the V_1 axis. The σV, V_1 plot for the same data is shown in fig. 2. It will be seen that the plot is linear on the left hand side of the diagram but the points move off the line as it approaches the V_1 axis. The slope of the straight part is 4·88, which corresponds to the temperature 2436° K., and the contact potential difference is 2·36 volts. This quantity, which gives the true position for zero field on the V_1 scale, is marked off on fig. 1. If equations (1)–(3) applied accurately, the

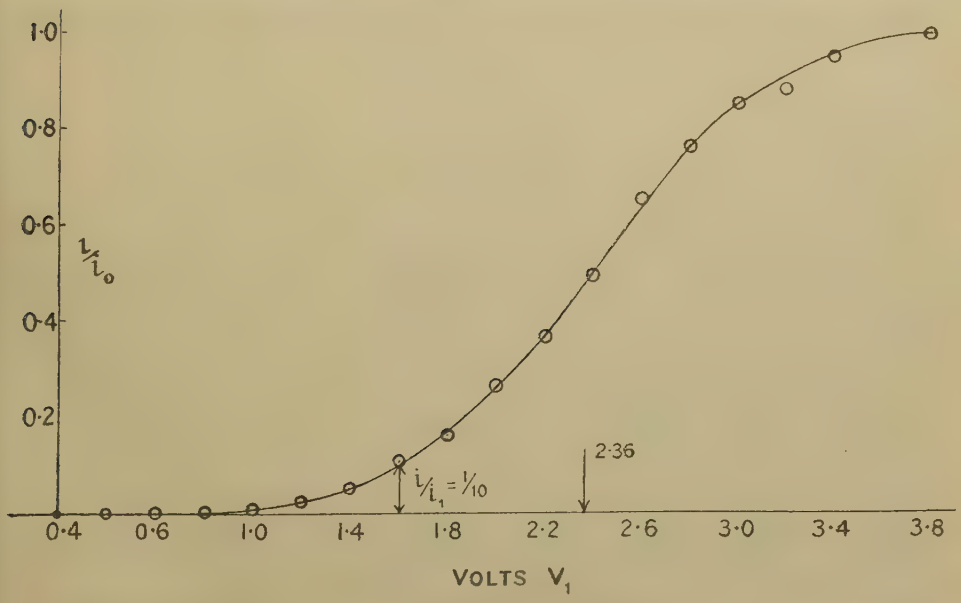


Fig. 1.

current should come to saturation at this voltage. As a matter of fact they only apply to the left of the point marked $i/i_0 = 1/10$. To the right of this point the currents do not rise fast enough as the retarding field falls, and in fact an appreciable accelerating field is required to bring the current to saturation. This deviation is the counterpart of the deviation of the points from the straight line on the right hand side of fig. 2.

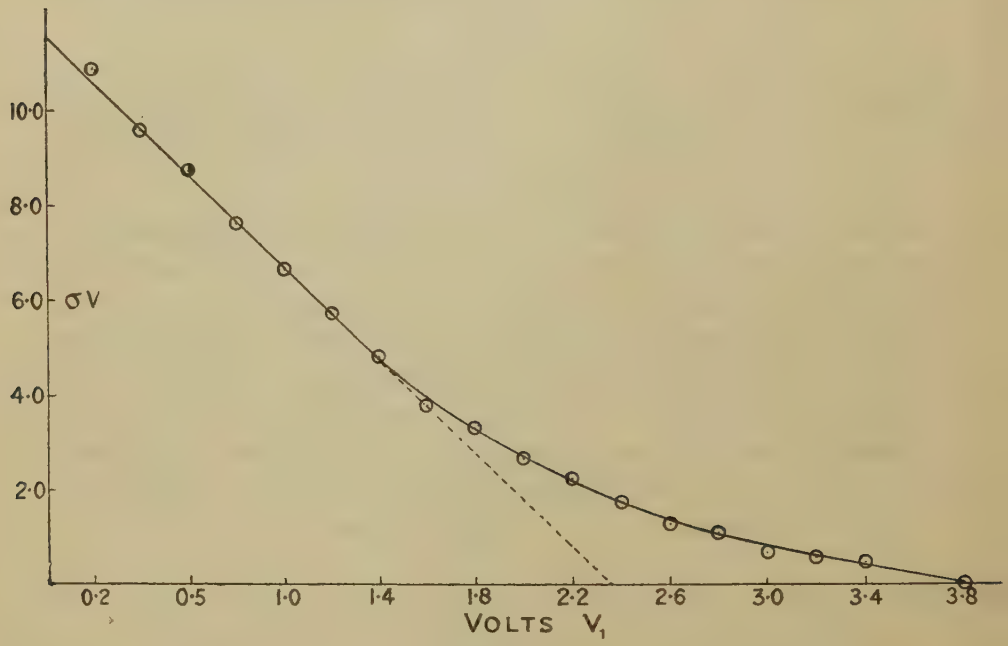


Fig. 2.

A selection of results from similar experiments taken under different sets of conditions is collected in Table V.

The compositions of the alloys which are indicated by the various designations Na K, Na K_2, the letters A–F and Nos. 1–4 in this Table and in various places in the rest of this paper are as follows :—Na K contained 63 per cent. of K, Na K_2 77 per cent. K, A not known with certainty but believed to be 59 per cent. K, B also not known with certainty but believed to be 50 per cent. K, C also not known with certainty but believed to be 50 per cent. K, D also not known with certainty but believed to be $>$ 59 per cent. K, E 53 per cent. K, F the same as Na K_2 77 per cent. K, No. 4 85 per cent. K, No. 1 $>$ 85 per cent. K, No. 2 the same approximately as No. 1, No. 3 more K than Nos. 1 and 2 but $<$ 90 per cent. K, No. 5 $>$ 85 per cent. $<$ 90 per cent. K.

Some idea of the fluidity of the alloys may be got from the following data.

Table V.

No. of Expt.	Compn. of Alloy.	Method of Exhaust.	Pressure mm. of Mercury.	Satn. Current i_0 (amps.).	Time 1 drop (secs.).	Contact P.D. Volts.	Slope.	Remarks.
1	Na K	Liquid Air	<0·001	$2·58\times10^{-9}$	17·7	−2·360	4·88	Brass sphere coated with silver—diam. 3·8 cms.
2	,,	,,	<0·001	$2·58\times10^{-9}$	17·7	2·42	4·88	,, ,,
3	$Na K_2$	Gaede	?	$9·6\times10^{-10}$	13·2	2·06	5·15	,, ,,
4	,,	,,	0·0065	3×10^{-10}	17·2	1·93	5·19	,, ,,
5	,,	,,	0·0065	$6·32\times10^{-10}$	16·2	2·02	5·17	,, ,,
6	,,	,,	0·0065	$3·45\times10^{-9}$	13·5	2·16	5·14	,, ,,
7	,,	,,	0·0065	$2·83\times10^{-9}$	13·5	2·22	5·06	,, ,,
8	,,	,,	0·0052	$3·66\times10^{-9}$	13·5	2·29	4·86	,, ,,
9	,,	,,	0·0065	$2·75\times10^{-9}$	13·5	—	5·22	,, ,,
10	?	,,	0·008	$2·8\times10^{-11}$	3·0	1·86	4·94	,, ,,
11	?	,,	0·0095	5×10^{-11}	3·1	1·96	5·23	,, ,,
12	?	,,	0·0070	$7·5\times10^{-11}$	3·03	2·09	5·13	,, ,,
13	?	,,	0·008	$1·2\times10^{-10}$	3·03	2·07	5·11	,, ,,
14	?	,,	0·0084	$1·5\times10^{-10}$	3·03	2·14	5·01	,, ,,
15	?	,,	0·0065	$1·13\times10^{-10}$	3·06	1·98	5·00	,, ,,
16	?	,,	0·0078	$2·2\times10^{-10}$	3·06	2·11	5·00	,, ,,
17	?	,,	0·0065	$2·93\times10^{-10}$	3·21	2·04	5·05	,, ,,
18	Na K	,,	?	5×10^{-9}	6·66	2·26	5·01	,, ,,
19	,,	,,	?	$7·3\times10^{-9}$	6·00	2·28	5·05	,, ,,
20	,,	,,	?	$9·16\times10^{-9}$	4·50	2·41	4·84	,, ,,
21	,,	,,	?	$9·16\times10^{-9}$	4·30	2·37	5·00	Silvered brass sphere, 3·8 cms.
22	,,	,,	?	$5·0\times10^{-9}$	8·0	2·37	4·95	,, ,,
23	,,	Liquid Air	<0·001	$2·5\times10^{-9}$	23·0	2·34	5·07	,, ,,
24	,,	,,	<0·001	$2·29\times10^{-9}$	22·0	2·28	4·99	,, ,,
25	,,	Gaede	0·014	$7·0\times10^{-9}$	7·1	2·39	5·80	,, ,,
26	,,	,,	0·021	$7·0\times10^{-9}$	4·6	2·45	6·10	,, ,,
27	,,	,,	0·029	$6·38\times10^{-9}$	7·2	2·54	6·38	,, ,,
28	B	,,	?	$5·6\times10^{-9}$	2·0	+0·11	4·87	Copper sphere, 1-inch diam.
29	B	,,	?	$2·6\times10^{-9}$	1·66	+0·05	4·75	,, ,,
30	B	,,	?	$3·33\times10^{-10}$	1·30	−1·90	5·10	,, ,,
31	A	,,	?	$2·27\times10^{-10}$	6·6	1·82	4·66	Silvered brass sphere, 3·8 cms.
32	A	,,	?	$1·3\times10^{-10}$	8·4	1·88	4·61	,, ,,
33	A	,,	?	$3·33\times10^{-10}$	10·0	1·24	4·00	,, ,,
34	B	,,	?	$7·16\times10^{-9}$	0·97	1·51	4·85	,, ,,
35	B	,,	?	$3·0\times10^{-10}$	9·0	1·84	4·93	,, ,,
36	B	,,	?	$2·5\times10^{-10}$	16·0	1·82	5·16	,, ,,
37	B	,,	?	$2·08\times10^{-10}$	17·5	1·86	4·80	,, ,,
38	B	,,	?	$4·58\times10^{-10}$	11·3	1·95	5·00	,, ,,
39	B	,,	?	$1·25\times10^{-10}$	12·0	1·98	5·00	,, ,,
40	D	,,	?	$5·8\times10^{-9}$	1·0	0·00	5·00	1-inch Cu. sphere.
41	D	,,	?	$9·0\times10^{-10}$	3·75	1·65	4·80	,,
42	$K/Na=\frac{53}{47}$	,,	?	$1·8\times10^{-9}$	?	2·44	5·00	,,
43	,,	,,	?	$2·3\times10^{-9}$	?	2·33	5·07	,,
44	$Na K_2$	,,	?	$1·33\times10^{-9}$	6·0	2·02	5·20	,,
45	,,	,,	?	$4·5\times10^{-9}$	?	1·49	5·10	,,
46	,,	,,	?	$2·8\times10^{-9}$	35·0	1·90	5·00	,,
47	No. 2	,,	?	$7·5\times10^{-8}$	1·15	2·34	4·99	,,
48	No. 2	,,	?	8×10^{-9}	11·3	2·28	5·10	,,
49	No. 3	,,	?	$1·37\times10^{-9}$	54	2·99	4·91	,,
50	No. 3	,,	?	8×10^{-10}	45	−3·23	5·03	,,

The most liquid alloy melts at — 13° C. and has the composition 77 per cent. K. This corresponds to the formula $Na K_2$ but it is not believed to be a compound. Those which melt below 0° C. contain from 67 to 85 per cent. K, and those which melt below + 15° C. contain between 47 and 90 per cent. K.*

As regards pressure of $COCl_2$ the experiments in Table V fall into 3 classes. In Nos. 1, 2, 23 and 24 the apparatus was shut off from the pump and the gases were absorbed by charcoal cooled by liquid air. This led to pressures as low as, or lower than, the limit of the Pirani gauge used (about 0·001 mm.). In Nos. 25, 26 and 27 a larger nozzle and higher gas pressures in the reservoir were used and the gases were removed by the Gaede mercury pump. In this way steady pressures above 0·010 mm. were obtained. In all the other experiments a finer nozzle and a lower gas pressure in the reservoir were used and the gases taken away by the Gaede pump. The pressures were only measured in Nos. 4 to 17, but it is believed that in all these experiments, which were carried out with very various objects, the pressures were within the limits 0·005 and 0·010 mm. of mercury.

The value of the contact potential difference for the surfaces in a normal condition probably varies with the composition of the alloy (and to a smaller extent, for the materials used, with the composition of the other electrode). It appears from the data to have a minimum value of about 2·0 volts for $Na K_2$ the most liquid alloy, and to rise to 2·35 volts for Na K and to 2·3 volts for alloy No. 2 which contained nearly 90 per cent. of potassium. The value for $Na K_2$ is in fair agreement with the results in (3) for $Na K_2$ and $COCl_2$ where the value of the contact P.D. was found to lie between the limits of 1·6 and 1·9 volts; at any rate both agree in giving low values with this alloy. The compositions given are those of the alloys when they were first made, and may not be altogether reliable as there are indications that the alloys become richer in Na after being subjected to the reaction. For this reason and on account of various factors which may cause the contact potential difference to be altered, too close a correspondence between the results of different experiments must not be expected. The data must be interpreted broadly.

In four of these experiments the outer electrode was observed to be splashed over with the alloy, with the result that in the three of them in which fast drops were used (Nos. 28, 29 and 40) the value of the contact P.D. was reduced practically to zero, and in the other (No. 33) it was low. This is satisfactory because for two surfaces each covered with the fresh alloy the contact P.D. should be zero. It is likely that most, if not all, the low values of K found in

* *Cf.* H. Le Chatelier, 'Receuil de Constantes Physiques,' Paris, 1913, p. 352.

Table V are the result of some splashing. The high values shown by Nos. 49 and 50 with an alloy rich in K may be due to a layer on the electrode (see §10 below).

It will be seen (§ 6) that there is a variation, with the composition of the alloy, in the threshold frequency which corresponds to this change in the contact potential. Experiments 24–27 were made expresssly to test the effect on the phenomena of increasing the pressure of the $COCl_2$. It appears that one effect is to increase the measured contact potential difference. It will be shown (§ 6) that the $COCl_2$ exerts an analogous effect on the threshold frequency, which it lowers. This corresponds to an increased electro-positiveness of the alloy and thus to a higher contact potential.

Perhaps the most remarkable thing in Table V is the constancy of the slope. If we exclude Nos. 25–27 and 31–33, of which we shall speak later, in none of the experiments does it differ as much as 5 per cent. from 5·00. This is in spite of the following variations in the independent variables :—Partial pressure of $COCl_2$ from $< 0{\cdot}001$ to $0{\cdot}010$ mm., saturation current (total emission) from $2{\cdot}8 \times 10^{-11}$ to $7{\cdot}5 \times 10^{-8}$ amp., time of one drop from 0·97 to 54 seconds, change in the composition of the alloy from approximately 50 per cent. K to nearly 90 per cent. K, and the substitution of a brass sphere 3·8 cms. in diameter coated with silver for a copper sphere 2·54 cms. in diameter.

Experiments Nos. 24–27 were made to test the effect of higher pressures of gas. These experiments should perhaps be repeated but they are consistent with one another in showing a steady increase in the measured values of the slope at pressures above 0·010 mm. of mercury. There are, of course, many causes which might give rise to such an increase in the slope, and it does not necessarily mean that there is any corresponding change in the energy of the emitted electrons under these higher pressures.

The reason for the low values given by Nos. 31–33 is not known. These also all gave low values of the contact potential. In the case of No. 33 splashing of the outer electrode was observed. It may be that the condition of the outer electrode was steadily changing in these experiments.

The data in Table V do not afford any clear evidence of any alteration of the slope with either (1) changes of pressure between $<0{\cdot}001$ and $0{\cdot}010$ mm., (2) total emission between $2{\cdot}8 \times 10^{-11}$ and $7{\cdot}5 \times 10^{-8}$ amp., (3) rate of drops between 1 per minute, and 60 per minute, and (4) composition of alloy between 50 per cent. K and 90 per cent. K.

The mean value of the 44 *consistent determinations of the slope*, i.e., *excluding Nos.* 25–27 *and* 31–33, *is* 5·014. *This corresponds to a value of* $T = 2368°$ *K.*

Some other experiments not included in Table V made with a cylindrical copper electrode 2·0 cms. in diameter, the one used in (3) in fact, gave practically the same value.

Fig. 3 shows the determination of σV as a function of V_1 obtained in 4 successive experiments, and is a fair specimen of kind of consistency which is obtained.

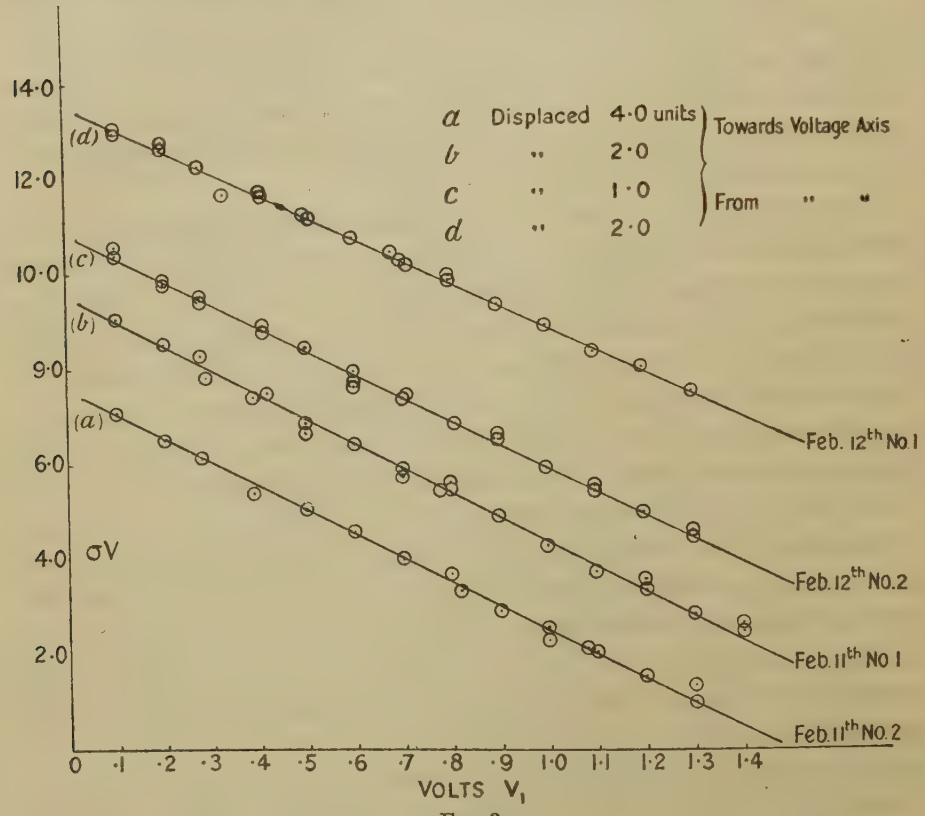


Fig. 3.

The successive sets of points are here displaced vertically by arbitrary amounts in order to separate them. It will be seen that the plots are accurately linear over a wide range, a range in fact which involves a reduction of the chemical current in the ratio 5000 : 1. This result is important because it establishes the asymptotic approach of the chemical current towards the voltage axis to the corresponding degree of accuracy. There is nothing in the chemical phenomenon which corresponds to the sharp intersection of the voltage axis which is represented by the photoelectric stopping potential.

§ 5. *The Photoelectric Determination of the Contact Potential.*

The difference of potential between the drops and the sphere was determined by photoelectric measurements in (3) where the method used is fully described. Briefly it involves the determination of the threshold frequency ν_0 at the drops and the stopping potential V_s for some definite frequency $\nu(> \nu_0)$, the contact potential K between the two surfaces then being given by

$$e\,(V_s + K) = \tfrac{1}{2}mv^2 = h\,(\nu - \nu_0). \qquad (5)$$

In this equation V_s is the voltmeter reading at which the photoelectric current is reduced to zero. In experiments Nos. 1–27 no photoelectric tests were made as we assumed that the identity of the contact potentials given by the chemical action method and by the photoelectric method was sufficiently well-established by the results in (3).

At this stage it became necessary to review the position as the slopes found in experiments Nos. 1–24, although consistent with each other, were uniformly and very appreciably higher than those measured in (3). It was felt that too much confidence was being reposed in the conclusion in (3) that the initial velocity distribution is invariably Maxwellian; particularly as the present experiments showed very wide deviations from such a distribution in the neighbourhood of zero velocity. It is not intended to suggest that these deviations are a serious argument *against* the phenomena involving a Maxwell distribution as a fundamental element. There are good reasons for believing that such deviations may arise from secondary causes. They cannot, however, be regarded as strengthening the case for a Maxwell distribution. It was decided to try to shed some more light on the phenomena by reverting to the photoelectric measurements.

As in (3) the source of monochromatic light consisted of various lines in the visible from a quartz mercury vapour lamp focussed on the drops by a Hilger glass monochromator. In order to improve the results we obtained a " K.B.B. microscope Illuminant " quartz mercury vapour lamp which is more powerful than the one used previously. There are considerable difficulties about these experiments. It is necessary to determine the photoelectric data under the same conditions as the chemical. This involves obtaining the photoelectric effect by subtracting the chemical effect (light off) from the sum of the chemical and photoelectric effects (light on). The chemical effect is in most cases so large that the result is zero unless a very powerful monochromatic source is used. The only lines, either from this source or any other, which we have been able to get in sufficient strength are the yellow group λ 5679–5790, the green line

λ 5460, the violet group λ 4347–4358 and the violet group λ 4046–4077 of mercury lines. The blue line λ 4916 is too weak to be useful and the intensity of the violet group λ 4046–77 seems to vary in relation to the intensities of the others in a way which we do not understand. At any rate we have found it very difficult to make any reliable determinations of the threshold frequency between $5 \cdot 5$ and about $6 \cdot 5 \times 10^{14}$. These difficulties are the cause of the gaps in the information presented in the next section. It is also difficult to determine the stopping potentials with an accuracy of more than one-twentieth of a volt. At the stopping potential the photoelectric current is necessarily zero, and in this neighbourhood the photoelectric currents are inevitably superposed on a much larger current of chemical origin.

There is little doubt that the assumption that the chemical and photoelectric effects are independent and superposable is true as a first approximation, but it is not certain that it is true as an exact principle. In some experiments with an alloy No. 2, which is rich in potassium, and with a slow rate of drops, conditions which give a large chemical emission accompanied by a relatively small photoelectric emission, it appeared that the currents with the light on were less than with the light off. This would imply that one effect of the light is to diminish the rate of chemical emission. This particular effect has, however, not been sufficiently investigated to warrant further discussion at this stage. It has also been observed with another alloy also containing more than 85 per cent. K but with a fast rate of drops (1 in $1 \cdot 66$ secs.).

The results of a particular experiment of the kind now under consideration are shown in fig. 4. This was made with alloy B. The points shown by $\odot$ represent the observations of the chemical effect, those shown by $\oplus$ the photoelectric currents with the violet group λ 4046–77, and those shown by $\times$ the currents with the violet group λ 4347–58. It will be seen that the zero on the V_1 scale given by the chemical emission is $1 \cdot 82$. In this experiment the value of ν_0 was close to $5 \cdot 0 \times 10^{14}$. When this is combined with the value of ν for the group λ 4347–58 we find $\frac{h}{e}(\nu - \nu_0) = 0 \cdot 77$ volts which gives the zero on the V_1 scale as photoelectrically determined $= 2 \cdot 46$ volts. Similarly the values from the group λ 4046–77 give the zero as $2 \cdot 49$ volts. These two photoelectric determinations are in good agreement with each other, but they are both very much higher than the figure given by the chemical determination.

A number of similar disagreements between the photoelectric and the chemical determination of the contact potential difference were obtained both with this and other alloys. In each case the chemical value of the contact potential

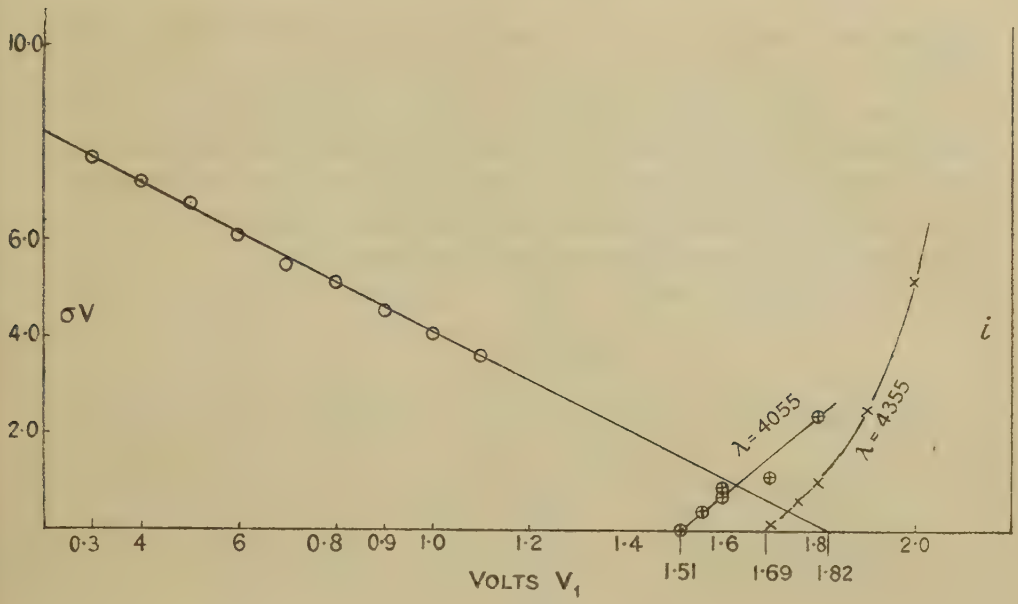


FIG. 4.

difference was less than the photoelectric. In order to clear up this discrepancy a more systematic investigation of both the chemical and photoelectric phenomena was undertaken with a number of alloys of known composition.

§ 6. *The Threshold Frequencies of Different Alloys.*

We have only been able to obtain satisfactory determinations of ν_0 both *in vacuo* and in presence of $COCl_2$ in the case of the alloys which contained a relatively low percentage of potassium. An alloy (marked E) which was made up of 53 parts K to 47 Na gave 3 values of ν_0 with $COCl_2$ present as $5{\cdot}05$, $5{\cdot}125$ and $5{\cdot}175 \times 10^{14}$ with a mean $5{\cdot}12 \times 10^{14}$. This point could be determined quite satisfactorily to the degree of consistency of these numbers because measurable effects were got with the green and yellow as well as with the violet lines. The vacuum value was determined as $> 5{\cdot}4 \times 10^{14}$. This point is less accurate but it is certainly not very far from the mark, and there is no doubt that it is higher than the value when $COCl_2$ is present and $< 6{\cdot}0 \times 10^{14}$.

Very close determinations of ν_0 for both conditions were got from another alloy after it had been used for several experiments. The value in $COCl_2$ was found to be $5{\cdot}0 \times 10^{14}$ and that *in vacuo* $5{\cdot}38 \times 10^{14}$. The composition of this alloy was unknown. Another alloy also of unknown but different composition gave $\nu_0 = 5{\cdot}40 \times 10^{14}$ in $COCl_2$ and $> 5{\cdot}40 \times 10^{14}$ *in vacuo*.

Still another experiment with alloy B gave $5{\cdot}0 \times 10^{14}$ in $COCl_2$ and $5{\cdot}2 \times 10^{14}$ *in vacuo*.

All three results agree in making the threshold frequency a little higher *in vacuo* than in $COCl_2$. Whether a similar result would hold with the alloys richer in K we are unable to state. They are, however, consistent with the observations in (3) where it was found that in experimenting in water vapour the threshold on one occasion was pushed into the infra-red.

The threshold frequencies for a number of alloys are assembled in

Table VI.

Designation of Alloy.	B.	E.	A.	F.	No. 4.	No. 1.	No. 3.	No. 5.
Percentage of K	?50	53	?59	77	85	>85	>No. 1	>85
$\nu_0 \div 10^{14}$ for alloy *in vacuo*	5·38	>5·40	>5·40	—	>5·40	—	5·05	5·45
$\nu_0 \div 10^{14}$ for alloy in $CO\,Cl_2$	5·0	5·12	5·40	6·2	5·20	5·09	—	—

$> 5{\cdot}4$ means that there was no emission with λ 5461 or that it was so small that it could not safely be accepted as real: it also implies $< 6{\cdot}0 \times 10^{14}$.

The queried compositions assigned to alloys A and B are only guesses but the others are certainties (within a certain degree of approximation). It is clear from the figures in the bottom row of Table VI that ν_0 in $COCl_2$ has a maximum at a composition near 77 per cent. K. This is the composition of the most liquid alloy which, although not believed to be a compound, corresponds to the formula $Na\,K_2$. It is probable that the vacuum value of ν_0 does not vary so much with the composition of the alloy as that measured in $COCl_2$.

The present observations are in agreement with the observations of the thresholds in (3) when account is taken of the composition of the alloy there used.

§ 7. *The Stopping Potentials for Different Alloys.*

The stopping potential for the photoelectrons ejected by the violet group λ 4046–77 has been determined for alloys containing 50 per cent. K or thereabouts, both with and without $COCl_2$ in the apparatus. For these alloys the stopping potential was found to be about $< 0{\cdot}1$ volt, or rather less, higher in $COCl_2$ than *in vacuo*. A particular determination of this difference, made with the alloy to which the first column of figures in Table VI refers and in the same state, gave the following figures:—Stopping potential with $COCl_2$ present $= 1{\cdot}668$ on the V_1 scale: *in vacuo* $1{\cdot}581$: difference $0{\cdot}087$ volt. In this experiment the pressure of the $COCl_2$ was about $0{\cdot}008$ mm. and the rate of drops 4 in 90 seconds.

Some data showing the relation between the threshold frequency ν_p, the

stopping potential V_s for light of frequency $\nu = 7{\cdot}40 \times 10^{14}$, and the chemically and photoelectrically determined contact potential differences for 6 alloys of different compositions are given in Table VII. The numbers given in each case are those found in the presence of $COCl_2$. The quantity D is the difference on the V_1 scale between the zero got from the chemical effect and the photoelectric stopping potential. The compositions of the alloys marked ? are estimated by interpolation from the threshold frequencies using the data in Table VI.

Table VII.

Designation of Alloy and per cent. of K.	No. of Expt.	$\nu_p \div 10^{14}$	$\frac{h}{e}(\nu-\nu_p)$ (volts).	Stopping Potential ν_s (volts).	K_p (volts).	K_c (volts).	K_p–K_c (volts).	D (volts).
B 50 ?	34	5·0	0·979	−1·07	−2·05	−1·51	0·54	0·44
B 50 ?	35	5·0	0·979	1·43	2·41	1·84	0·57	0·41
B 50 ?	36	5·0	0·979	1·48	2·46	1·83	0·63	0·35
B 50 ?	37	5·0	0·979	1·51	2·49	1·82	0·67	0·31
B 50 ?	38	5·0	0·979	1·63	2·61	1·95	0·66	0·32
B 50 ?	39	5·0	0·979	1·65	2·63	1·98	0·65	0·33
E 53	42	5·11	0·943	1·98	2·92	2·44	0·48	0·46
E 53	43	5·11	0·943	1·97	2·91	2·33	0·58	0·36
D>59 ?	40	>5·4	<0·816	+0·40	<0·416	0·00	<0·42	0·40
D>59 ?	41	>5·4	<0·816	1·25	<2·066	1·65	<0·41	0·40
F. 77	44	6·2	0·489	1·62	2·11	2·02	0·09	0·40
F. 77	45	6·2	0·489	1·04	1·529	1·49	0·039	0·45
F. 77	46	6·2	0·489	1·40	1·889	1·90	+0·001	0·50
No. 2>85	47	5·05	0·957	?1·46	?2·42	2·34	−0·08	?0·88
No. 2>85	?48	5·06	?0·95	1·82*	?2·75	2·28	0·47	0·36*
No. 3>No. 2	?49	4·65	?0·91	2·39*	?3·30	2·99	0·31	0·51*
No. 3>No. 2	?50	4·65	?0·91	2·83*	?3·74	3·23	0·51	0·31*

* It was not possible to measure the stopping potential in $COCl_2$ in these experiments. To get the values of D, 0·09 volts was added to the vacuum values which are given in the fourth column. This is the correction which has to be applied to alloys poor in K, but it is not known whether it is applicable to these alloys.

The numbers in the second column relate these experiments with the data for the same experiments in Table V. The third column contains the photoelectric threshold frequency ν_p. K_p (col. 5) is the contact P.D. as measured photoelectrically and deduced from the equation

$$K_p = \frac{h}{e}(\nu - \nu_p) - V_s, \qquad (6)$$

K_c is the chemical contact potential got as described in § 4. It will be seen that these quantities are only equal for the most liquid alloy, that which contains 77 per cent. K. In general $-K_p > -K_c$, the divergence being greater the more the composition deviates from 77 per cent. K either in the direction of more K or of more Na. Although the individual values of K_c or of K_p vary

by large amounts in a fortuitous manner, the difference $K_p - K_c$ is a quite systematic function of the composition of the alloy to the accuracy of the experiments. (It should be remembered that the data got from the alloys Nos. 2 and 3 containing $>$ 85 per cent. K are much less reliable than the others.) The fortuitous variation in K_p and K_c is due to changes in the condition of the outer electrode in different experiments. This is eliminated when we take the difference $K_p - K_c$. In the experiment which gave $K_c = 0{\cdot}00$ the outer electrode was observed to be splashed over with the alloy.

The change in $K_p - K_c$ with the composition of the alloy is the volt equivalent of the corresponding change in the photoelectric threshold frequency ν_p, to the accuracy of the data. The difference between K_c and K_p is evidently closely associated with the changes in ν_p. It is probably due to some peculiarity in K_p as the quantity D, which is the photoelectrically determined stopping potential measured from the chemically determined zero, shows relatively little variation, if any, with composition. This point will be reconsidered in the following section after some other matters have been discussed.

§ 8. *Comparison with the Results of* (3).

The results of §§ 6 and 7 dispose of one of the conflicts between the results of the present investigation and those described in (3). The particular conflict is, that whereas in (3) the chemically and photoelectrically determined contact potentials were found to be identical for the same surfaces, it appeared in the present experiments, § (5), that there might be a very considerable difference between them. It appears from Table VII that in the present experiments the two contact potentials are identical for the alloy which contains 77 per cent. K. This is the most liquid alloy having a composition corresponding to Na K_2 and is the alloy which was used in the only experiments in (3) for which an accurate determination of the contact potential was made by both methods.

The results are, however, still not entirely harmonious. In the old experiments the average energy of the electrons, expressed as an equivalent temperature, was found to be 3300° K. within a probable error thought to be about 10 per cent. The corresponding quantity found in the present investigation is 2368° K. as the mean of 44 determinations, none of which differ from this value by so much as 5 per cent.

Whether the difference between $T \simeq 3300°$ K. in (3) and $T = 2368°$ K. in the present experiments represents anything real must still be left in suspense. It is hoped to settle this question definitely by further experiments. The difference is about 30 per cent., and it hardly seems probable that the former

experiments can involve so large an error. The conditions were definitely different in the two sets of experiments. In (3) the emission was varying with the pressure of the $COCl_2$, as is shown by the response to the liquid air control, whereas in the present experiments the emission was independent of the pressure at any rate over most of the range used (§ 3). The natural interpretation of this fact is that we are now dealing with the average energy of the electrons emitted up to the completion of the reaction over a given metal surface ; whereas at the lower pressures which appear to have prevailed in (3) only the incipient stages of this reaction may have been concerned. The chemical currents in (3) were also smaller than in the present experiments. If there is a real difference in the two cases it should be in the direction indicated by the figures which have been obtained, as the average electron energy would be cut down by the necessity of getting through the surface layer of reaction products in the later stages of the reaction.

The present results tend to strengthen the arguments in favour of the view that the phenomenon is a direct chemical emission and not due to secondary thermionic effects (*cf.* (2), p. 313). If the old experiments are reliable, as it seems likely they are, the measured energies were about 30 per cent. above those now found. In any event they are certain enough to show that there is no considerable increase in the energy when the chemical currents are increased from the smallest values which can be measured to the highest values attainable by raising the gas pressure. The highest chemical currents used (in one set of experiments or the other) are roughly $3{\cdot}7 \times 10^4$ times the smallest. This constancy is incompatible with the view that the phenomena are secondary thermionic effects. This argument is, however, not so conclusive when regarded from the standpoint of the theory of patches (8.9).

§ 9. *The Difference between the Photoelectric and the Chemical Zero.*

The difference between the apparent contact potential as determined from the properties of the chemically emitted electrons and by the photoelectric measurements (§ (7)) is a matter of considerable interest and calls for further discussion. It can be accounted for if it is assumed that in general the chemical reaction does not take place uniformly over the surface of the drops, but occurs in localised patches which grow by extension of the periphery. There is already definite photoelectric and thermionic evidence of the occurrence of patches having different properties from the major part of the surface in the case of the alkali metals.* If we admit the existence of such patches in the case of the

* O. W. Richardson and A. F. A. Young, ' Roy. Soc. Proc.,' A, vol. 107, p. 377 (1925).

chemical reactions here investigated we already have information which enables us to say something about their photoelectric properties. If we refer to Table VI we find that in each case in which it was possible to determine ν_0 both with and without $COCl_2$ the value was a little higher without $COCl_2$. This means that the surface acted on by the $COCl_2$ is electropositive to the clean alloy. This may be surprising but it fits with the observations of Richardson and Young and also with the very low thresholds for the alloy which have been found in the presence of water vapour. Let ν_p denote the threshold frequency for the contaminated areas and ν_c that for the clean alloy. The photoelectric measurements will always determine the smaller of these and for the cases where both have been ascertained this is ν_p. On the view taken it is no longer surprising that the contact potential is different according to whether it is determined from the chemical or from the photoelectrons as these come from surfaces with different intrinsic potentials. The stopping potential V_s of the photoelectrons is given by the equation

$$V_s = \frac{h}{e}(\nu - \nu_p) - K_p.$$

There is a relation between K_c and K_p namely

$$K_c = K_p + \frac{h}{e}(\nu_p - \nu_c).$$

If we make use of this relation we can find an expression for D the quantity in the last column of Table VII. We have

$$-D = K_c + V_s = \frac{h}{e}(\nu - \nu_c).$$

It is evident from Table VII that D changes very little with the composition of the alloy, nothing like so much, for example, as $\frac{h}{e}(\nu - \nu_p)$ which is given in the third column. Evidently ν_c is much less dependent on the composition of the alloy than ν_p. This is also in agreement with the data in Table VI where the vacuum data should correspond to ν_c and the $COCl_2$ data to ν_p. It is hoped that it may be possible to make further experiments to test some of these relations more accurately.

The theory of patches accounts for another set of facts very satisfactorily, and that is the difficulty of attaining saturation, and the breaking down of the agreement with the Maxwell formula in the neighbourhood of zero emission velocity. The experimental conflict with the formula is of this character that the proportion of low velocity electrons is far below what it ought to be.

If the electrons are ejected at the periphery of an electropositive patch the slow ones will be pulled into the positively charged area and only the fast ones will get away. It is found in addition, in accordance with the same ideas, that an accelerating field is required to draw a considerable fraction of the slow electrons into the surrounding spherical surface. These phenomena are not found with the photoelectric currents which approach the saturation axis much more steeply.

§ 10. *Hysteretic Effects.*

These effects are partially described in (4) where the values of the currents at a given potential difference vary according to whether the potentials are rising or falling. They probably account for most of the discrepancy between the results described in (1) and those in (2), (3) and (4), and also for most of the troublesome shifts along the voltage axis which occur from time to time. Our understanding of these effects, while still imperfect, has been greatly increased as the result of experiments with alloy No. 4, composition 85 per cent. K, which, for some unknown reason, exhibited the effects in a very exaggerated form. The phenomena can be best explained by reference to fig. 5. The

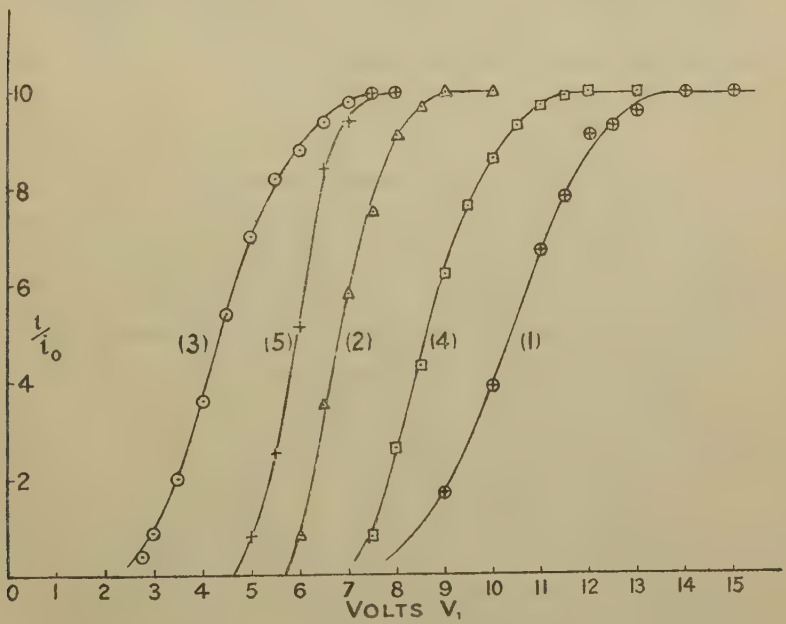


Fig. 5.

chemical current E.M.F. curve with points :—⊕ was first taken, 15 volts being applied to saturate the current between each reading. The gas was then

pumped out and the photoelectric curve for the line λ 4055, points :—△, was taken. The apparatus was allowed to stand 18 hours, $COCl_2$ re-admitted and the chemical curve given by ⊙ obtained without applying 15 volts between readings. The chemical curve got by applying 15 volts between readings was then taken and found to be that given by ⊡. The gas was then taken out and the photoelectric response to λ 4055 was then found to be given by +. This is not a complete statement ; several other things were done during these experiments, but it is believed, as the result of additional experiments, that the changes which caused the observed effects have been mentioned.

The essential feature of these and of innumerable other hysteresis phenomena which we have examined is that the application to the outer electrode of a positive potential which is fairly high, but not too high, causes the currents subsequently measured at lower potentials to occur at higher potentials than would otherwise have been the case. The phenomenon is, in fact, similar to the polarisation of a voltaic cell. This cranky or polarised state may, as in the example shown in the figure, last for hours, or it may be a matter of minutes or seconds, or the time of relaxation may be so quick that the effects appear to be absent. In addition to the displacements along the voltage axis shown in fig. 5 the individual curves differ from a normal curve, such as that in fig. 1, by being drawn out horizontally.

The site of this effect is at the surface of the outer electrode. It is quite definitely not at the surface of the drops because (1) it can persist for hours, whereas the surface of the drops is renewed several times a minute, and (2) whereas it may move the photoelectric stopping potential for λ 4055 from about 3 to about 6 volts there is no corresponding change in the photoelectric *threshold frequency*. The last fact alone shows that the nature of the surface of the drops is not seriously altered when the effect is present. It is not in the gas because the displacement of the photoelectric stopping potential for λ 4055 persists when the gas is pumped out and the tests made in a vacuum. For the same reason it cannot be a temporary effect caused by the action of the gas on the surface of the outer electrode. It must be a change in the surface of this electrode which is capable of being fairly permanent.

Effects of this character have been observed over the whole range of pressure of $COCl_2$ covered by these experiments down to $<0{\cdot}001$ mm. of mercury, with slow and with fast drops, and with alloys covering nearly the whole range of composition which is liquid. The most marked effects were obtained with an alloy containing a very high percentage of K, but they were not noticeable in experiments with an alloy containing a still higher percentage of K. The

effects have been observed with surfaces of copper, brass and silver, and with surfaces which may be new, old, dirty or clean. When the condition is very pronounced it appears that it may be destroyed by raising the potential difference, driving the electrons from the drops to the outer electrode, to 100 volts or less. As a result of this property the hysteresis curves embracing higher potentials of this order may be extremely complicated.

We are glad to acknowledge our indebtedness to the Government Grant Committee of the Royal Society, and to the Department of Scientific and Industrial Research for financial assistance in connection with this investigation.

Gaseous Combustion at High Pressures.—Part VII.* *A Spectrographic Investigation of the Ultra-Violet Radiation from Carbonic Oxide-Oxygen (or -Air) Explosions.*

By William A. Bone, D.Sc., F.R.S., and D. M. Newitt, Ph.D.

(Received March 31, 1927.)

[Plate 1.]

Introduction.

The radiation emitted by flames in various circumstances was not much studied until the year 1890 when R. von Helmholtz and (independently) W. H. Julius made the first systematic analyses of the quantity and quality of that emitted by the flames of hydrogen, carbonic oxide, methane and ethylene burning in air at ordinary pressures. Their work, however, does not seem to have been much noticed until, in 1907, H. L. Callendar directed attention to it in connection with the work of the British Association Committee on Gaseous Explosions, which was chiefly concerned with how far radiation is responsible for what is sometimes termed "the missing pressure" in closed vessel explosions.

In 1910 B. Hopkinson (who was assisted by W. T. David) published his measurements of the total radiation emitted during a coal-gas-air explosion

* Previous papers of this series appeared in 'Phil. Trans.,' A, vol. 215, p. 275 (1915); 'Roy. Soc. Proc.,' A, vol. 100, p. 67 (1921); *ibid.*, vol. 103, p. 205 (1923); *ibid.*, vol. 105, p. 406 (1924); *ibid.*, vol. 108, p. 393 (1925); *ibid.*, vol. 110, p. 645 (1926).

at an initial pressure of one atmosphere in a closed cylindrical vessel (30 cm. × 30 cm.) and the subsequent cooling period. The explosive mixture employed was one of 15 parts coal-gas with 85 of air, such being very nearly the mixture of maximum strength consistent with complete combustion. Two comparative sets of experiments were made, namely :—(i) with the walls of the vessel highly polished, and (ii) in which they were painted with a 0·02 mm. layer of a dead-black lamp-black mixed with a little shellac. The explosion chamber was fitted with a fluorite window, outside of which was fixed a resistance-bolometer, made of a blackened platinum strip, for measuring the radiation emitted.

The results showed that the *total* energy radiated during and after the explosion amounted to about 25 per cent. of that represented by the *gross* calorific value of the gas burnt; howbeit, only one-seventh of such total radiation was emitted during the actual explosion, *i.e.*, up to the attainment of the maximum pressure. Subsequently W. T. David showed that in all probability the radiation emitted in such explosions mainly consists of the same two bands in the infra-red (4·4 μ and 2·8 μ respectively) as are emitted by an ordinary Bunsen flame. Of these the 4·4 μ band is due to incipiently formed (or forming) CO_2 molecules, and the 2·8 μ band to incipiently formed (or forming) H_2O molecules.

Very little is definitely known as yet concerning the influence of temperature and pressure upon the radiation emitted during gaseous explosions. The British Association Gas Explosions Committee were content merely with saying that "within moderate limits of pressure, the radiating and absorbing power of a flame per unit thickness at a given temperature and composition should vary directly as the pressure or density." And, in regard to temperature, they remarked that, although Nernst had deduced that the radiation varies as the fourth power of the temperature, "his conclusion was most severely criticised by Lummer, Pringsheim and Schaefer, who explained that the radiations were quite different from that of a black body, and that the quality of the radiation was little, if at all, affected by pressure up to four atmospheres." Also they pointed out that, according to Planck's formula for a single wave-length, the rate of variation "is much lower than the fourth power law, and tends in the limit to be directly proportional to the absolute temperature at high temperatures," and therefore, "the actual rate of variation should lie between these limits, but nearer to Planck, *unless carbon begins to separate in rich mixtures at high temperatures.*"

As long ago as 1858 E. Frankland observed that in general the luminosity

of flames increases with pressure, and subsequently Liveing and Dewar connected the phenomenon with an increase in the intensity of the continuous spectrum with pressure, which they showed is proportional to the square of the pressure. They inferred that "the brightness of a continuous spectrum depends mainly on the mutual action of the molecules of the gas," and added that "in every case the prominent feature of the light emitted by flames at high pressures appears to be a strong continuous spectrum. There is not the slightest indication that this continuous spectrum is produced by the widening of the lines, or obliteration of the inequalities of discontinuous spectra of the same gas at lower pressures."

There is already a good deal of evidence that radiation functions more in gaseous explosions than merely as dissipating a considerable part of the energy liberated; indeed, it may well be that it plays an important rôle in *initiating* the chemical changes commenced by "activating" the participating gases.

During the researches which we have carried out in recent years at the Imperial College, London, in conjunction with Dr. D. T. A. Townend, results have been obtained on exploding CO-air mixtures at high initial pressures which seem best explained on the supposition that, in such circumstances, nitrogen absorbs part of the radiation emitted by the burning carbonic oxide, being thereby "activated" and rendered capable of combining with any excess oxygen present forming nitric oxide.

The outstanding facts, in connection with such explosions are:—(i) that, *contrary to all other cases yet examined*, progressive increases in the initial pressure between 10 and 175 atmospheres caused a progressively increasing "time-lag" in the attainment of the maximum pressure; (ii) that as soon as such "lag" first became manifest, and for a considerable range of initial pressure thereafter (10 to 25 atmospheres), it was always accompanied by a marked exothermic effect during the "cooling period" without any appreciable lowering of the "corrected" P_m value, and (iii) that when more oxygen was present than that required to burn all the carbonic oxide, a secondary NO-formation occurred which was favourably influenced by increasing pressures, and at high initial pressures it considerably exceeded that corresponding with a merely thermal equilibrium in the system $N_2 + O_2 \rightleftharpoons 2NO$ at the maximum explosion temperature.

All these facts, and some others connected with CO-air explosions under pressure, may be explained on the supposition (i) that in a CO-oxygen explosion the radiation emitted is absorbed by the combustible gas which is thereby "activated" so as to be capable of combining directly with the oxygen,

(ii) that when nitrogen is present as a "diluent," as in a CO-air explosion, part of the radiation emitted is intercepted and absorbed by the nitrogen so that the rate of activation of the CO molecules is retarded, producing the observed "lags" in the attainment of maximum pressure, and (iii) that the nitrogen so "activated" *either* (*a*) in absence of excess oxygen, reverts gradually to the ordinary state during the "cooling period," liberating as heat the equivalent of radiant energy which it absorbed during the explosion period, or (*b*) when excess oxygen is present, combines gradually with it forming nitric oxide.

According to this view, such NO-formation is a *secondary* phenomenon, dependent upon a previous N_2-activation. It presumably requires an appreciable time, and therefore may not always begin, or be recognisable, during the actual "explosion period." In such cases it would occur wholly during the succeeding "cooling period." On the other hand, at high initial pressures it may begin during the explosion period, *i.e.*, *before* the attainment of maximum pressure, and continue into the "cooling period." And, from the results of our previous bomb experiments at various initial pressures,* it was concluded that under such conditions no appreciable secondary NO-formation would begin during the explosion period in a CO-air explosion unless the initial pressure exceeded 25 atmospheres or thereabouts.

Having arrived at such conclusions from experiments in which the observations made were chiefly pressure-time records supplemented by chemical analyses, it was thought desirable to carry the matter a step further by seeing what a spectrographic analysis of CO-air, etc., explosions might reveal. This meant the designing of an explosion chamber with quartz windows specially adapted for the spectrographic examination of explosion flames. Thanks to generous aid from the Advisory Council of the Department of Scientific and Industrial Research the necessary financial provision was forthcoming, and we were able to install the new apparatus and begin the experiments about two years ago, since which time the work has been continuously in progress. The present paper contains a first instalment of our results relating to $2CO + O_2 + xR_2$ explosions (where $R_2 = O_2$, CO or N_2) at initial pressures between 10 and 25 atmospheres.

Apparatus.

(1) *The New Cylindrical Bomb No. 4.*

The new cylindrical bomb used in our spectrographic experiments is shown diagrammatically in fig. 1. It was made out of a nickel-steel forging by Messrs.

* 'Roy. Soc. Proc.,' A, vol. 105, p. 406 (1924).

Armstrong Whitworth & Co., Ltd., of Newcastle-on-Tyne, and we desire now to thank Sir Albert G. Hadcock, F.R.S., their Managing Director, for his kind

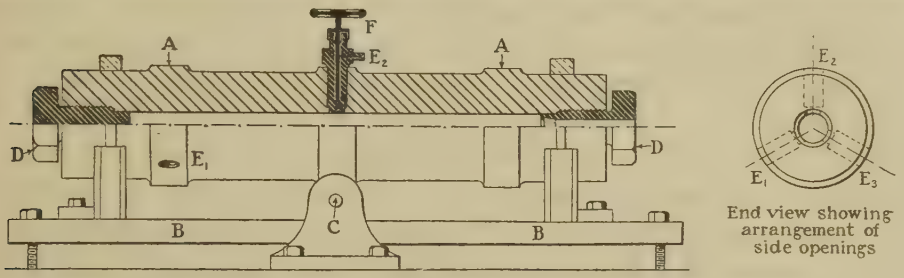


FIG. 1.

help and advice in connection with its design. Its principal dimensions are :— overall length, 1 metre ; external diameter, 20 cms. ; internal length of explosion chamber, 80 cms. ; diameter of explosion chamber, 5 cms. ; capacity of explosion chamber, 2 litres.

The bomb AA is mounted horizontally on a cast iron stand BB pivoted at the centre C so that it may be rocked. It is closed at the ends by the steel plugs DD. There are two pairs of such plugs, one pair carrying quartz windows, and the other pair being blank.

The way in which the quartz windows A are fitted into the plugs is shown in fig. 2. Each window is formed out of a perfectly clear unstrained circular piece of quartz 2·5 cms. long, ground with a slight taper so as to fit accurately into a similarly tapered opening of the plug, the joint being made tight by a steel collar B and asbestos packing. It was found that windows so fitted are capable of withstanding the shock of explosions in which pressures of up to 400 atmospheres are suddenly generated, and that the joints so made between quartz and steel are perfectly gas-tight under such conditions.

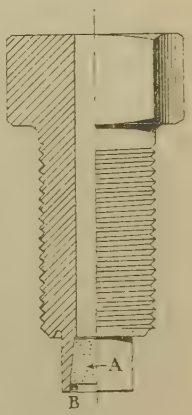


FIG. 2.

In addition to such plugs and windows for screwing into each end of the explosion chamber, there are also (*a*) an inlet valve, and (*b*) a special pressure gauge of the Petavel type, either of which can be substituted for them as and when required.

Along the barrel of the bomb are three side openings $E_1E_2E_3$—one in the middle, and one about 10 cms. from each end—into which can be screwed either an inlet valve, or an ignition plug, or a subsidiary quartz window plug, or a blank plug, according to experimental requirements.

In making-up and firing an explosive mixture in the bomb, each constituent is separately introduced, through a P_2O_5 drying-tube, by means of the admission valve F, perfect mixing being subsequently ensured by rolling a 1·5 cm. diameter steel ball many times to and fro along the whole length of the explosion chamber. Ignition is effected by the fusion of a thin platinum wire in the firing plug.

(2) *The Spectrograph.*

A Hilger E_2-quartz spectrograph, fitted with a Cornu prism, and having a dispersion of 20 A.U. per mm. at 3110, has been used throughout the experiments. In carrying out an experiment, the slit of the spectrograph was placed close to the quartz window in one of the end plugs of the explosion chamber of the bomb, no intervening lenses or diaphragms being necessary. For reference purposes an iron-arc spectrum was usually photographed on each plate, the arc being placed in front of a similar quartz window fitted into the other end of the explosion chamber so that no change was made in the relative positions of the bomb and spectrograph. In cases where the spectra of two successive explosions were being photographed on one plate, a Hartman diaphragm was used, thus obviating the necessity of moving the plate between the two explosions. The exact parallelism of the jaws of the slit of the spectrograph was verified at the outset of the experiments.

Experimental.

(1) *On the Formation of Oxides of Nitrogen in CO-Excess-Air Explosions at High Initial Pressures.*

The object of these experiments was to ascertain spectrographically whether or not the marked formation of nitric oxide observed when a mixture of carbonic oxide with excess of air is exploded at an initial pressure of (say) 25 atmospheres occurs during the actual "explosion period," *i.e.*, before the attainment of maximum pressure.

For this purpose a mixture of two volumes of carbonic oxide with $6\frac{1}{4}$ volumes of air was exploded in the bomb at an initial pressure of 25 atmospheres; the arrangement of the apparatus was as shown in fig. 3, the explosion-chamber

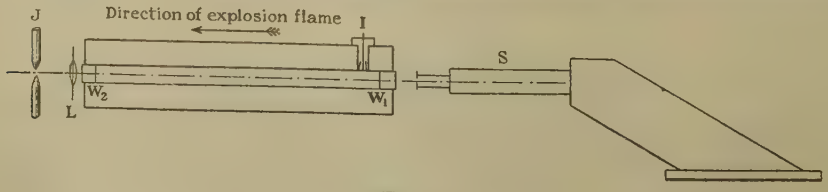


Fig. 6.

being fitted with a quartz window (W_1W_2) at each end. The explosive mixture was fired at the end of the explosion chamber nearest the spectrograph S and window W_1 by the fusing of a thin platinum wire in the ignition plug I; the resulting spectrogram would therefore record the radiation from the explosion after it had traversed the hot but gradually cooling products.

On observing the explosion through the window W_2, a bright orange-coloured flame could be seen travelling along the explosion chamber followed immediately by a red glow which persisted for some time after the flame had ceased. On afterwards examining the cooled products by transmitted light they were seen to have a deep red colour, and subsequent analysis showed them to contain 2·8 per cent. of nitrogen peroxide.

Immediately *before* and *after* exploding the mixture in question, spectrograms (Plate 1, A and C) were taken of an ordinary CO-flame burning in air from a jet placed at J in front of window W_2, so that the radiation from it was converted by the quartz lens L into a parallel beam which then traversed *either* (*a*) the explosive mixture at 25 atmospheres pressure *or* (*b*) the explosion products reduced to 1 atmosphere pressure before reaching the spectrograph as the case might be. In this way, the three spectrograms A, B, C were all obtained on the same plate, namely:—

Spectrogram *A*, of the radiation from a flame of *undried* CO burning in air at ordinary pressure after it had traversed the cold unexploded CO-air mixture at 25 atmospheres pressure in the explosion chamber of the bomb. This shows the familiar continuous spectrum between 4400 and 2300 A.U. together with steam lines in the region of 3089 A.U.

Spectrogram *B* of the radiation from the $2CO + 6\frac{1}{2}$ air explosion in the bomb at an initial pressure of 25 atmospheres. This shows the continuous CO-flame spectrum between 4400 and 2800 A.U., but neither "steam lines" nor any sign of NO-lines or NO_2 bands.

Spectrogram *C* of the radiation of a flame of *undried* CO burning in air at ordinary pressure after it had traversed the cooled explosion products whose pressure had been reduced to 1 atmosphere by opening the valve F.* It shows the characteristic NO_2-absorption bands overlying the continuous CO-flame spectrum, as well as some "steam lines" in the region 3089 A.U.

* It was necessary to blow off the surplus pressure of the gaseous products in the bomb at this juncture, otherwise no appreciable radiation from the CO-flame would have been transmitted through them owing to their high NO_2-content.

The results of this experiment prove that, in the given circumstances, (*a*) the CO in the mixture exploded was burnt without the *intervention of steam*, (*b*) no appreciable NO was formed during the "explosion period," or indeed whilst radiation of wave-length shorter than 4400 A.U. was being emitted, although (*c*) NO_2 was abundantly produced at some later period.

In order to gain some idea of the mean maximum temperature attained in the foregoing explosion, another experiment was made with the same mixture at the same initial pressure; but instead of taking a spectrogram as in the former case, a pressure-time record was taken by means of a modified Petavel gauge, as follows:—

Initial Pressure = 25 atmospheres
Maximum Pressure = 160 atmospheres
$t_m = 0{\cdot}45$ sec.
Mean Maximum Temperature = 2120° A.

According to Nernst's calculation, the proportion of NO in thermal equilibrium with a system $3{\cdot}8\ N_2 + O_2$ (air) at 2120° A would be 0·8 per cent. only, whereas in our experiment no less than 2·8 per cent. of it was found in the cold products. This difference is all the more remarkable when it is considered that in our experiment almost four-fifths of the oxygen in the air originally present had been consumed by the carbonic oxide before any NO-formation began.

(2) *Explosions of* $CO + N_2O$ *and* $CO + N_2O + NO$ *mixtures.*

In order to be quite sure that if any NO had been formed during the "explosion period" in the foregoing experiment, evidence of it would have been shown in the spectrogram, we next took comparative spectrograms of the explosions of mixtures corresponding with

(*a*) $CO + N_2O$, and (*b*) $CO + N_2O + NO$

when fired in the bomb at initial pressures of 6 and 9 atmospheres respectively. These are reproduced in Plate 1, D and E respectively, both explosions being extremely violent almost to "detonation."

It will be seen that D shows an intense continuous CO-flame spectrum, on which appear two absorption lines at 3248 and 3274 A.U. respectively. These are undoubtedly due to copper, traces of which may occasionally get into the bomb from the long narrow-bore copper tubes of the filling system. In the case of E, where the mixture exploded initially contained NO, a series of regularly distributed absorption bands is superposed on the continuous CO-flame spectrum.

(*Facing p.* 48.)

For purposes of comparison, we reproduce in Plate 1, F, the emission spectrum obtained when nitric oxide is excited in a " vacuum tube " by a high tension alternating current; it is identical with that given by either nitrous oxide or nitrogen containing traces of oxygen in similar circumstances. It will be seen that it differs both in wave-lengths and their distribution from the absorption bands shown in E. Prof. A. Fowler, F.R.S., has been unable to identify the latter with any known banded-spectrum. And inasmuch as in our experiments these unknown bands have been observed only when nitric oxide has been present in the CO-mixture exploded, but never in either a CO-air or a $CO + N_2O$ explosion at high initial pressures, they would seem to be due to nitric oxide under the conditions of the extremely high temperatures and pressures generated in our explosions.

Measurements of the bands in question are tabulated in Table I. Owing to the fact that many of the bands are faint, and all are broad and ill-defined (possibly due to pressure), the wave-lengths are subject to a probable error of 3 to 5 A.U. The intensities are estimated on a scale of 3 for the brightest bands to 0 for bands of which a trace only is visible on the negative. The distribution of the bands is regular; considering the approximate character of the wave-numbers, the analysis given in Table II indicates that they form part of a definite banded spectrum of molecular origin.

Table I.

Wave-length.	Intensity.	Wave number.
3835	1	26,077
3720	0	26,881
3680	3	27,173
3640	0	27,472
3580	0	27,931
3545	3	28,209
3500	2	28,569
3440	0	29,070
3405	3	29,369
3370	2	29,674
3290	0	30,395
3255	1	30,720
3220	1	31,055

Table II.

31055	(1)	*1381*	29674	(2)	*1105*	28569	(2)	*1097*	27472	(0)
335			*305*			*360*			*299*	
30720	(1)	*1351*	*29369*	(3)	*1160*	28209	(3)	*1036*	27173	(3)
325			*299*			*278*			*292*	
30395	(0)	*1325*	29070	(0)	*1139*	27931	(0)	*1050*	26881	(0)

3.—*On the Absorption by Nitrogen or Carbonic Oxide of the Ultra-Violet Radiation from a* $2CO + O_2$ *Explosion under Pressure.*

Evidence that either nitrogen or carbonic oxide will strongly absorb the radiation emitted during a $2CO + O_2$ explosion under pressure was forthcoming from the experiments described in this section of our paper, in which photometric measurements were made of the comparative densities on the same plate of the spectrograms from explosions of $2CO + O_2 + 5R_2$ mixtures (where $R_2 = O_2$, CO, or N_2—*i.e.*, all diatomic diluents) all at an initial pressure of 13·7 atmospheres, the results of which will now be detailed.

Experimental procedure.—Briefly the experimental procedure was (i) to obtain a spectrogram from each of the three explosions in question ($R_2 = O_2$, CO or N_2 respectively) on one and the same plate, and afterwards (ii) to measure their relative densities at certain selected wave-lengths.

The actual explosions were carried out in the bomb as already described except that they were fired from the end remote from the spectrograph so that in each case the radiation emitted by the flame would have to traverse the unburnt mixture in front of it before reaching the spectrograph. As a matter of fact, however, the results of some comparative experiments, under the same conditions except that the explosion mixtures were fired at the other end (*i.e.*, nearest to W_1 and the spectrograph), showed that it was immaterial from which of the two ends firing took place. Three successive explosions were made with (i) $2CO + O_2 + 5O_2$, (ii) $2CO + O_2 + 5CO$, and (iii) $2CO + O_2 + 5N_2$ mixtures, respectively, at the same initial pressure, the resulting spectrograms being taken on one and the same plate. Between each successive explosion both of the quartz windows were removed and very carefully cleaned before being replaced in the bomb.

We would here point out that in the explosions referred to always the same absolute quantities of carbonic oxide and oxygen combined at the same initial pressure, in the same bomb, yielding products containing the same proportion of carbon dioxide and a diatomic diluent, the densities and thermal capacities of the three media being almost equal.

In such circumstances it seems reasonable to suppose that the radiation originally emitted by the combining molecules during combustion would be qualitatively and quantitatively the same in each case, and, therefore, that any differences in the relative densities of the resulting spectrograms would be due, not to differences in emission, but to differences in the subsequent absorption by the medium of the radiation originating during each explosion.

Measurement of relative spectrogram-densities.—The densities of the spectrograms were measured by means of a comparison photometer fitted with a standard optical wedge and photo-electric cell, the density of the image being given on the instrument directly by the scale reading.*

The general principles involved in comparing intensities by photographic methods have been fully discussed in a recent communication to 'Nature'† by Dr. Toy, with whom we conferred about the matter. Where conditions permit, the photographic plate is used solely as a detector of equal intensities of the same kind of radiation, the only assumption being that "on adjacent portions of a plate, radiation of the same intensity, of the same quality, acting for the same time, will produce the *same* effect when the plate is developed."

When the intensity and quality of the radiation and the time of exposure vary, a knowledge of the characteristics of the photographic plate is necessary; this is obtained by a calibration of the plate using radiations of known wave lengths.

In experiments in which the radiation to be measured originates from the explosive combustion of gaseous mixtures of different composition, both the intensity of radiation emitted and the time taken for complete combustion may vary. It should be noted, however, that in the mixtures investigated by us these quantities are of the same order, a fact which considerably facilitated calibration.

It should also be mentioned that, since in all the experiments now to be described the same amount of combustible gas was burnt under approximately the same pressure conditions, the quantity of radiation emitted would presumably be constant, and that consequently the product of the intensity of total radiation and time of combustion (exposure) would also be constant.

The calibration of our plates therefore included a careful determination of the value of "p" (the so-called Schwarzschild's constant), in the expression $D = It^p$ for radiations of wave-length 3400, 3800 and 4200 A.U. The plates used in all intensity determinations were W. and W. Rapid 250 H. and D., and we were able to satisfy ourselves that the value of "p" for these plates remained substantially constant for the range of intensities and times of exposure obtaining in our experiments. The constancy of "p" under these conditions has been shown by the work of Jones and Huse‡ who recorded a

* The instrument used was that employed by the British Photographic Research Association, and we wish to express our thanks to Dr. T. Slater Price and to the Association for placing the instrument at our disposal and for advice in making the measurements.

† 'Nature,' January 16, 1926.

‡ 'J. Opt. Soc. Am.,' vol. 2, p. 1923 (1925).

maximum variation of 3 per cent. only for a typical rapid plate over the same range of intensities.

In comparing the intensities of radiation emitted during the explosions of the different gaseous mixtures, the respective spectrograms were always taken on the same plate in neighbouring positions near the centre and within short intervals of one another, thus eliminating any errors due to development or to variations in the quality of the emulsion. No attempt has been made to compare spectrograms not on the same plate, although the conditions of development have been kept as nearly constant as possible throughout the research.

In fig. 4 density measurements for the complete ultra-violet spectrum of a typical carbon monoxide-air explosion at an initial pressure of 25 atmospheres

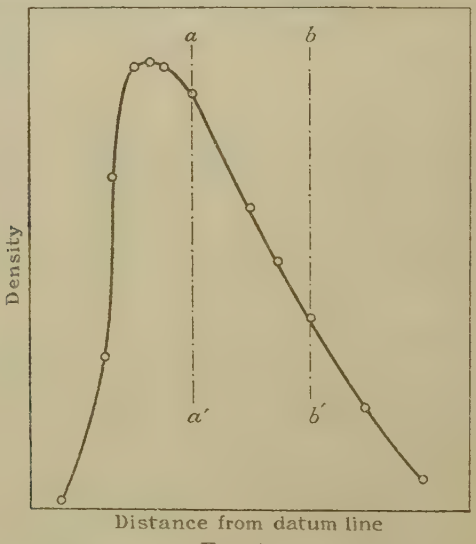


Fig. 4.

are plotted against distance from a given datum line on the negative. The rapid increase of density to a maximum in the neighbourhood of 5000 A.U. is observed in all such curves, and is due to the increasing sensitiveness of the plate with diminishing wave length in this region. The true maximum for the explosion would be at about 4·6 μ in the infra red. Over the region of uniform sensitiveness of the photographic plate the density of the spectrum shows a gradual falling off, and at 2500 A.U. it is almost negligible. All measurements of intensity have been made from that portion of the density curve lying between the two vertical lines aa' and bb' (*i.e.*, between 4400 A.U. and 3200 A.U. approximately).

Experimental results.

(1) *With $2CO + O_2 + 5R_2$ Mixtures.*—The character of the results obtained in these comparative experiments will be understood from Table III and the curves shown in fig. 5.

Table III.—Comparative Intensities of Radiations resulting from Explosions of Mixtures $2CO + O_2 + 5R_2$, where $R = N_2$, CO and O_2, respectively.

Wave-length A.U.	$2CO + O_2$ diluted with—		
	$5O_2$.	5CO.	$5N_2$.
4400	100	58·0	53·0
4300	98·5	57·8	52·8
4200	87·2	57·8	52·2
4100	80·0	56·5	51·2
4000	75·8	55·8	50·8
3900	70·2	54·0	49·0
3800	65·0	50·9	47·6
3700	63·3	49·5	46·8
Temp. °C.	2020	2220	1865

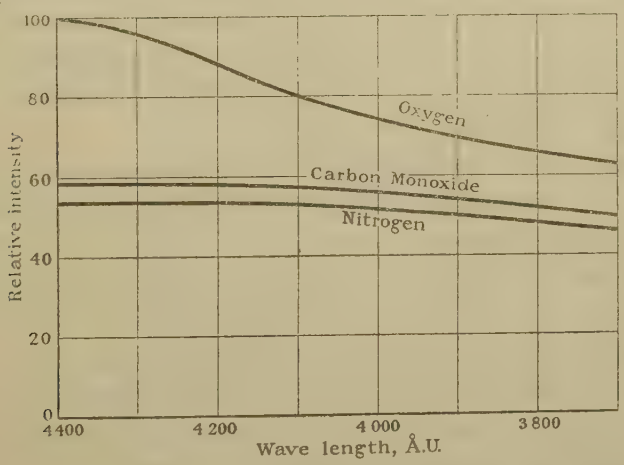


Fig. 5.

Table III shows the relative intensities of the radiation at various wave lengths between 4400 and 3700 A.U. resulting from explosions of the three $2CO + O_2 + 5R$ mixtures (where $R = O_2$, CO, or N_2, respectively) at an initial pressure of 13·7 atmospheres in each case. As a basis of comparison the intensity of the 4400 A.U. radiation from the explosion of the

$2CO + O_2 + 5O_2$ mixture is taken as 100, all other intensities being expressed as percentages of the standard. An estimate of the maximum temperatures attained in the three explosions in question is included in the Table.

It will be seen, both from Table III and the curves shown in fig. 5, that the intensity of the radiation resulting from the $2CO + O_2 + 5O_2$ explosion was, at all the wave lengths compared, very much greater than for either of the other two explosions where the " diluent " was carbonic oxide or nitrogen. This marked difference cannot be attributed to any difference in flame-temperature, because from Table III it will be seen that the maximum temperature attained in the $2CO + O_2 + 5CO$ explosion was about 200° C. higher than in the $2CO + O_2 + 5O_2$, which in turn was higher than that attained in the $2CO + O_2 + 5N_2$ explosion.

Whilst there is nothing in the experimental data to show how much (if any) of the ultra-violet radiation from the $2CO + O_2$ explosion was absorbed by the $5O_2$-diluent in the first experiment, it is difficult to resist the conclusion that a considerable part of such radiation was absorbed by the 5CO- and the $5N_2$-diluents in the other two experiments.

(2) *With* $2CO + O_2 + xR_2$ *Mixtures.*—In view of the foregoing conclusion it seemed of interest now to compare the intensities of the radiation resulting from $2CO + O_2$ explosions when similarly diluted with progressively increasing proportions of oxygen, carbonic oxide and nitrogen respectively. For this purpose explosions of mixtures of the type $2CO + O_2 + xR_2$, when $x = 4$, 5 or $5\frac{1}{2}$ and $R_2 = O_2$, CO or N_2, were studied.

The partial pressure of the $(2CO + O_2)$ detonating mixture in all experiments was 5·14 atmospheres, so that the $2CO + O_2 + 4R_2$ mixtures were all fired at an initial pressure of 12 atmospheres, the $2CO + O_2 + 5R_2$ mixtures at 13·7 atmospheres, and the $2CO + O_2 + 5\frac{1}{2}R_2$ mixtures at 14·6 atmospheres, respectively.

The results of the experiments are detailed in Table IV. It should be understood, however, that inasmuch as three different photographic plates were used, namely, one for the oxygen-series, another for the CO-series, and a third for the N_2-series, the results are not strictly comparable as between the three different series, but only within one and the same series. The basis of comparison in each series was the intensity of the radiation at 4400 A.U. emitted during the $2CO + O_2 + 4O_2$ explosion of that series. These experiments bring out very well the marked difference between the relative diminutions in radiation intensity, at all wave lengths examined, resulting from increasing O_2-dilution on the one hand, and from similarly increasing CO or N_2 dilution

on the other. There was no great difference, however, between the effects of the two latter, although at all except some of the shortest wave lengths a given dilution with nitrogen had a greater effect than the same dilution with carbonic oxide.

Table IV.—Comparative Intensities of Radiation resulting from Explosions of $2CO + O_2 + xR_2$ Mixtures where $R = O_2$, CO or N_2, and $x = 4$, 5 and $5\frac{1}{2}$, respectively.

Wave-length A.U.	$2CO + O_2$ diluted with—								
	Oxygen.			Carbon Monoxide.			Nitrogen.		
	$x = 4$.	$x = 5$.	$x = 5\frac{1}{2}$.	$x = 4$.	$x = 5$.	$x = 5\frac{1}{2}$.	$x = 4$.	$x = 5$.	$x = 5\frac{1}{2}$.
4400	100	69·0	69·4	100	39·5	36·2	100	36·0	30·0
4300	95·2	64·2	63·2	94·0	39·4	36·1	99·5	35·8	29·9
4200	88·9	60·0	57·8	87·5	39·1	35·9	99·0	35·5	29·8
4100	82·8	55·9	53·2	81·0	38·6	35·3	95·0	35·0	29·5
4000	76·8	52·2	48·6	74·2	38·0	34·8	90·0	34·2	29·3
3900	70·7	49·0	45·4	67·5	37·0	33·8	82·5	33·5	28·9
3800	65·4	46·0	42·6	60·0	35·5	32·7	76·0	32·5	28·4
3700	60·5	43·8	40·3	52·5	33·6	31·4	69·5	32·0	27·5
3600	56·0	41·6	38·1	46·2	32·0	30·0	65·0	31·4	26·5
3500	51·7	39·6	36·5	41·0	30·5	29·0	56·0	30·7	25·6
3400	47·6	38·1	35·6	36·5	29·4	28·0	50·5	30·3	25·0
3300	44·7	37·2	33·9	33·0	28·0	27·4	44·5	30·0	24·5
3200	41·6	36·6	33·0	30·0	27·5	26·7	39·0	29·5	24·1

(4) *Comparison of Ultra-Violet Radiation-Intensities from* $2CO + O_2 + 4N_2$ *and* $2CO + O_2 + 5CO$ *Explosions both developing the same Maximum Temperature.*

In the course of the investigation it was found that if a $2CO + O_2$ detonating mixture at 5·14 atmospheres were diluted with either $4N_2$ or 5CO, respectively, and the resulting $2CO + O_2 + 4N_2$ and $2CO + O_2 + 5CO$ mixtures exploded in the bomb at initial pressures of 12·0 and 13·7 atmospheres, respectively, so that precisely the same amount of carbonic oxide and oxygen combined in each case, the same maximum flame temperature (estimated to be as nearly as possible 2500° A) was attained in each explosion. Therefore it seemed of interest to compare the intensities of the ultra-violet radiation resulting from the two explosions. Accordingly the experiment was carried out with the results shown in Table V, the basis of comparison being the intensity of the radiation at 4400 A.U. emitted during the $2CO + O_2 + 4N_2$ explosion.

Table V.—Comparison of Intensities of Radiations from Explosions of Mixtures $2CO + O_2 + 4N_2$ and $2CO + O_2 + 5CO$, respectively, both developing the same Maximum Temperature.

Wave-length A.U.	$2CO + O_2$ diluted with—	
	$4N_2$.	5CO.
4400	100	38·1
4200	96·2	36·5
4000	91·5	35·6
3900	80·1	33·5
3800	73·0	32·8
3700	64·5	32·6
3600	57·4	31·7
3500	50·2	31·3
3400	43·2	30·4
3300	37·6	29·2
3200	35·0	26·5

Judging from these results it does not seem as though maximum temperature had much to do with the resulting ultra-violet radiation intensity, but rather that it was determined by the relative absorptive capacities of the $4N_2$ and 5CO diluents, the latter (partial pressure = 8·6 atms. when cold) absorbing more than the former (partial pressure = 6·9 atms. when cold).

We think that this experiment shows very well the strongly absorbing power which carbonic oxide has for the ultra-violet radiation from its own combustion. And it may be added that this feature was also shown in another pair of experiments in which the ultra-violet radiation resulting from $2CO + O_2 + 4CO$ and $2CO + O_2 + 6CO$ explosions at initial pressures of 12·0 and 15·4 atmospheres respectively (*i.e.*, *partial pressure* of the $2CO + O_2 = 5·14$ atms. in each case) were compared on the same plate. The first named explosion (*i.e.*, of the $2CO + O_2 + 4CO$ mixture) gave a spectrogram of considerable density between 4400 and 2900 A.U., but the radiation resulting from the second explosion (*i.e.*, of the $2CO + O_2 + 6CO$ mixture) was hardly strong enough to produce any image on the photographic plate.

(5) *The Effects of Monatomic Diluents upon the Ultra-Violet Radiations from $2CO + O_2$ Explosions at High Pressures.*

We have also compared the intensities of the resultant ultra-violet radiation from $2CO + O_2 + 4Ar$ or $2CO + O_2 + 4He$ explosions with that for a $2CO + O_2 + 4N_2$ explosion, all at the same initial pressure of 14 atmospheres. The results are summarised in Table VI, the standard of comparison

being the intensity at 4200 A.U. of the resultant radiations from a $2CO + O_2 + 4Ar$ explosion.

It should be here mentioned that the explosions were much more violent and the maximum temperatures attained much higher with the argon- and helium-diluted mixtures than when nitrogen was the diluent. The maximum temperature attained would be approximately 2790° in the Ar-diluted, 2660° in the He-diluted, and 2230° Cent. in the N_2-diluted explosion; in the first two cases, however, the percentage CO_2 dissociation at the maximum temperature would be about 15, but in the last case about 5 per cent. only.

From the comparative results shown in Table VI, it will be observed that the intensity of the resultant radiation from the $2CO + O_2 + 4Ar$ explosion was decidedly greater than that for the $2CO + O_2 + 4He$ explosion over the whole ultra-violet range. It thus seems as though argon has a specific influence upon the emission and/or absorption of radiation during a $2CO + O_2$ explosion which an equivalent amount of helium either does not exert at all or only in a much lesser degree. Also, as was anticipated from our previous experiments with CO-air mixtures, the resultant radiation from the $2CO + O_2 + 4N_2$ explosion was much feebler than even of that from the $2CO + O_2 + 4He$ explosion at the same initial pressure.

Table VI.—Comparative Intensities of Radiations from Explosions of Carbon Monoxide-Oxygen ($2CO + O_2$) Mixtures, diluted with Argon, Helium, and Nitrogen respectively.

Wave-lengths A.U.	$2CO + O_2$ diluted with—		
	4Ar.	4He.	$4N_2$.
4200	100	35·4	12·8
4000	83	31·6	11·7
3800	61·7	26·3	9·8
3600	44·8	25·5	8·4
3400	31·4	21·7	6·5
3200	22·5	16·4	5·1
Temp. °C.	2790	2660	2230

Summary.

The experiments described herein have shown that:—

(1) The resultant ultra-violet radiations from $2CO + O_2 + 4R$ explosions at corresponding high initial pressures, where R = a diatomic diluent, is much less when the latter is carbonic oxide or nitrogen than when it is oxygen;

this result indicates that carbonic oxide or nitrogen strongly absorbs the ultra-violet radiation emitted by the burning carbonic oxide in such circumstances.

(2) The absorption referred to is general, though not uniform, over the ultra-violet range A.U. 4400 to 3200°; and although there are no absorption bands or lines, the absorption is proportionately stronger at lower than at higher frequencies.

(3) The marked NO-formation which always occurs in a CO-excess-air explosion at an initial pressure of 25 atmospheres does not take place during the actual combustion, or indeed until after all the resulting radiation capable of affecting a sensitive photographic plate has been emitted.

(4) When NO is present during the actual combustion period in a CO-explosion, a definite absorption band spectrum is superposed upon the characteristic continuous ultra-violet spectrum of the burning carbonic oxide.

(5) The resultant ultra-violet radiation from a $2CO + O_2 + 4Ar$ explosion at an initial pressure of 14 atmospheres is very much stronger than that from a $2CO + O_2 + 4He$ explosion at the same pressure, although the maximum temperatures attained in the two cases differ by 130° C. only.

The results referred to in (1), (2) or (3), whilst generally supporting the view that either carbonic oxide *or* nitrogen is "activated" by the radiation emitted by burning carbonic oxide at high pressures, indicate that in the case of nitrogen such "activation" is of a lesser degree than that involving the formation of the "active" form of nitrogen discovered by the present Lord Rayleigh. Also, the results of the investigation generally show that the intensity of the ultra-violet radiation from the explosions in question depends on the maximum temperatures only in a subordinate degree.

We propose extending these spectrographic investigations to a study of the infra-red radiation emitted during such explosions as are described herein, and in other directions. Also, we have designed a new bomb which will enable us to undertake a systematic study of gaseous explosions at initial pressures between 200 and 1000 atmospheres, which will form the next section of our programme.

In conclusion we desire to thank the Department of Scientific and Industrial Research for grants out of which the cost of the bomb and spectrographic apparatus was defrayed, and which have also enabled one of us (D.M.N.) to devote his whole time to the research; moreover, we are indebted to the Government Grant Committee of the Royal Society for a grant out of which part of the incidental expenses of the work has been defrayed.

An Investigation of the Rate of Growth of Crystals in Different Directions.

By Miss Marie Bentivoglio, D.Phil. (Oxon) (1851 Exhibition Scholar).

(Communicated by Sir Henry Miers, F.R.S.—Received May 7, 1926. Revised* February 3, 1927.)

I.—Introduction.

The present investigation of the rate of growth of crystals in different directions was originally suggested by a paper by G. Wulff† on the velocity of growth and dissolution of crystal faces.

Wulff carried out an extensive series of experiments with "Mohr's Salt," iron-ammonium sulphate, $Fe(NH_4)_2(SO_4)_2 . 6H_2O$, growing a crust of this salt on a nucleus of the less soluble isomorphous zinc-ammonium sulphate, $Zn(NH_4)_2(SO_4)_2 . 6H_2O$, which was introduced into a saturated solution of the former salt. He measured the thickness of the crust deposited on various crystal faces, and found it to vary for different forms. As the thickness of the crust deposited represented the amount of growth of each form, Wulff deduced numerical values for the velocity of growth of the forms developed, taking as unity the rate of growth of r $\{\bar{2}01\}$, the slowest growing form observed.

Wulff then sought to establish a relation between the rates of growth of the various forms and their reticular densities, and on the assumption that the space lattice was of the simple (pseudo-) cubic type came to the conclusion that a low rate of growth corresponded, for the most part, to a high reticular density and vice versa. He pointed out that this conclusion agreed with the well-known fact that the faces most commonly occurring on a crystal are those with simple indices and high reticular density; since rapidly growing faces (if formed at all) tend to become smaller owing to the extension of surrounding faces, and finally to disappear.

The present paper is in great part concerned with an extension of Wulff's experimental work to other double sulphates. Experiments have also been made to determine whether the initial habit has any influence on the rate of

* Some of the tables giving the detailed results of the experiments, which were originally included in the paper, are here omitted, the mean results only being given. Owing to the absence of the author in Australia the revision involved has been done by Prof. H. L. Bowman.

† 'Z. f. Kryst.,' vol. 34, p. 449 (1901).

growth of the various faces. The latter part of the paper deals with the question whether a polar crystal may grow at different rates at its two ends—a question which does not seem to have been examined experimentally, though unequal rates of growth have been observed in microscopic crystals of triphenylmethane.*

It was hoped to discuss the results in relation to the reticular density of the faces; but until the actual space-lattices of the salts in question have been established by means of X-rays, such a discussion was felt to be premature.

One of the chief difficulties hitherto encountered in the determination of relative rates of growth has been that of obtaining uniform results. The initial object to be attained was, therefore, an improvement of the experimental methods, and a method was finally evolved which is not only more generally applicable than that employed by Wulff, but is also more accurate and gives more constant values. It has also the advantage that the same crystal can be used repeatedly for a series of experiments, and this facilitates the study of changes in habit during growth.

In considering the increase in size of a crystal, one must take into account, as Wulff has shown, the two aspects of growth: firstly, as an advance of each face along its normal from its point of origin, and secondly, as a tangential growth resulting in the enlargement of the faces. This tangential growth ("*Ausbreitung*") of a face may be either positive or negative, and is dependent on the relative rates of the "normal" growth of that face and of its neighbours, and on the angles between them. In the present research the growth has always been determined along the normals, and the effect of this upon the relative sizes of the faces, *i.e.* on the habit of the crystal, has been obtained by construction.

II.—Experimental.

1. *Material Used.*

Preliminary trials were carried out with octahedra of potash alum and cubes of sodium chlorate, to test the accuracy of the various experimental methods employed, but the substances selected for the research proper were the isomorphous double sulphates of magnesium-ammonium, iron-ammonium and magnesium-potassium, of the general formula $R''R'_2(SO_4)_2 . 6H_2O$, belonging to the monoclinic system, the second mentioned being that which formed the basis of Wulff's investigation. These salts were chosen as having suitable solubilities, and as being typical of this series of double sulphates and differing

* H. B. Hartley and N. Garrod Thomas, 'Trans. Chem. Soc.,' vol. 89, p. 1019 (1906).

considerably in average habit. Moreover, each salt commonly crystallises in two or three habits depending on the face on which the crystal rests during the early stages of growth; and crystals varying in habit—though exhibiting the same forms—were expressly used to determine the relation between habit and rate of growth.

The axes to which these crystals are referred are as follows:—

	a	: b :	c	β
Magnesium-ammonium sulphate	0·7400	: 1 :	0·4918	107° 6′
Iron-ammonium sulphate	0·7466	: 1 :	0·4950	106° 48′
Magnesium-potassium sulphate	0·7413	: 1 :	0·4993	104° 48′

The crystals of magnesium-ammonium sulphate were either stoutly prismatic in habit, sometimes with a well-developed face of $b\{010\}$ as shown in fig. 1, or tabular parallel to a face of $p\{110\}$ with small faces of $r\{\bar{2}01\}$ and well marked faces of $q\{011\}$ (fig. 2).

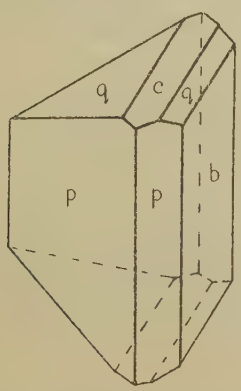


Fig. 1.

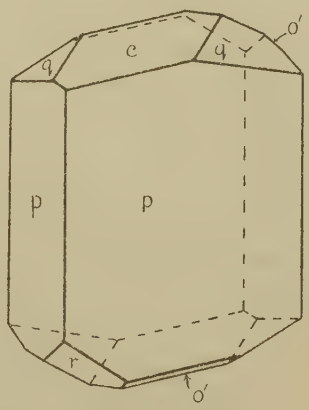


Fig. 2.

Magnesium-ammonium sulphate.

The most characteristic habit assumed by Mohr's salt was that of flattened crystals, tabular parallel to the faces of $r\{\bar{2}01\}$ (fig. 3), while a few crystals

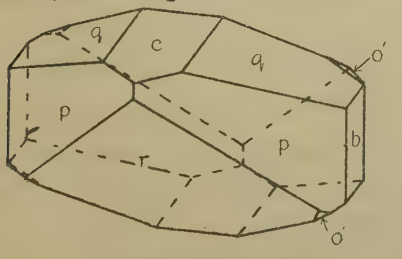


Fig. 3.

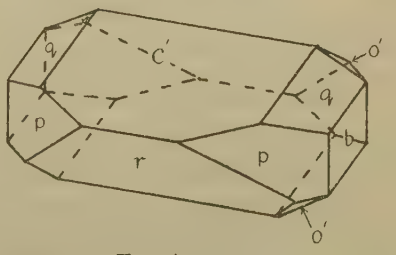


Fig. 4.

Iron-ammonium sulphate.

were tabular parallel to $c\{001\}$ (fig. 4), but r remained a predominant form and the prism zone was always weakly developed.

Crystals of magnesium-potassium sulphate were usually prismatic; faces of $a\{100\}$ were often developed, those of $r\{\bar{2}01\}$ and $q\{011\}$ when present were very small (fig. 5). The form $o'\{\bar{1}11\}$, which was small in the other salts, was almost always absent here.

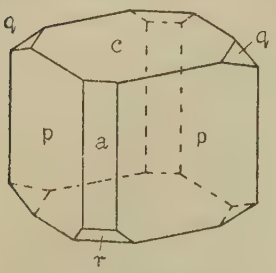


Fig. 5.—Magnesium-potassium sulphate.

The form $o\{111\}$ was occasionally seen in crystals of Mohr's salt, but it was very small and disappeared quickly with growth, and it was never measured. Rarely a small face of $p'\{120\}$ was seen in crystals of magnesium-potassium sulphate, but this quickly diminished and disappeared on growth under the conditions of the experiments.

For the experiments on polar crystals, ammonium tartrate, $(NH_4)_2C_4H_4O_6$, and potassium tartrate, $K_2C_4H_4O_6 \cdot \frac{1}{2}H_2O$, which both crystallise in the polar class of the monoclinic system, were chosen.

The axial constants to which these substances are referred are as follows:—

Ammonium tartrate .... $a : b : c = 1{\cdot}1506 : 1 : 1{\cdot}4383$, $\beta = 92° 23'$
Potassium tartrate .... $a : b : c = 3{\cdot}0869 : 1 : 3{\cdot}9701$, $\beta = 90° 50'$

The crystals were well formed and tabular in habit parallel to $a\{100\}$. The faces of interest here were those at the two ends of the axis of polarity. In the potassium salt (fig. 6) the face $b'(0\bar{1}0)$ truncates the negative end of the crystal,

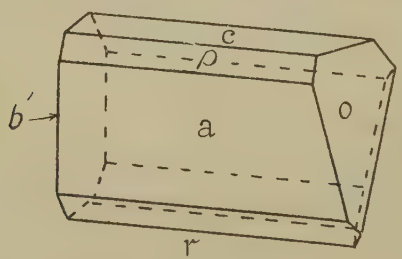


Fig. 6.—Potassium tartrate.

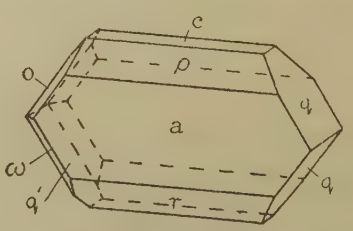


Fig. 7.—Ammonium tartrate.

and the sphenoid $o\{111\}$ forms the positive end. In the ammonium salt (fig. 7) the sphenoids $q\{011\}$ and $q'\{0\bar{1}1\}$ are the most prominent terminations, but small faces of $o\{1\bar{1}1\}$ and $\omega\{1\bar{1}\bar{1}\}$ also occur, generally at one end only. Other forms present in both substances were $c\{001\}$, $\rho\{101\}$, $r\{10\bar{1}\}$.

All the salts employed were purified by repeated recrystallisation.

The crystals selected for the experiments were small and generally trans-

parent. But crystals of magnesium-ammonium sulphate were cloudy even when they were very slowly formed. The faster growing faces showed greater cloudiness, but this did not affect the regularity of growth. When examined under the microscope, the cloudiness was seen to be due to the inclusion of rows of minute cavities filled with solution.

Crystals for measurement were obtained from slightly super-saturated solutions by slow cooling, or by evaporation at the ordinary temperature in a desiccator. They generally formed on the bottom of the glass vessel, so that the face on which they rested became somewhat concave. Before measuring any crystal it was therefore found expedient, after mounting it, to grow it under conditions similar to those of the ordinary experiments till all the faces showed even surfaces.

The crystals were smaller than those used as nuclei by Wulff. Two typical crystals measured about $3 \times 3 \times 4$ mm. and $2 \times 3 \times 5$ mm. at the beginning of a set of experiments. After a total growth lasting for 72 hours and 65 hours respectively, these crystals had grown to $15 \times 13 \times 16$ mm. and $10 \times 12 \times 15$ mm. Crystals of Mohr's salt grew especially well, and one crystal reached a size of $28 \times 30 \times 23$ mm. If, after a series of experiments, a transparent crystal was held up to the light, the original "nucleus" could be seen surrounded by several distinct layers, marked by inclusions, each layer representing the thickness of substance deposited during one experiment. It was interesting to note the change in habit of a crystal during growth, by comparing the nucleus and the final form.

2. *Apparatus and Experimental Methods.*

As the value of the present kind of research obviously depends on its accuracy, much time was devoted to designing appropriate forms of apparatus, and determining the conditions for obtaining the most trustworthy and uniform results.

The experimental work fell naturally into two parts, the actual growth of a crystal and its measurement. The former involved the careful control of conditions affecting growth, the latter required an accurate method of measuring the thickness of the deposited layers—preferably without destroying the crystal.

In Wulff's earlier experiments, in which the crystals were grown in a restricted quantity of solution, without stirring, the results were unsatisfactory, and Wulff was led to recognise that crystals grown in a fixed position do not grow regularly—an observation which has been confirmed in the present work.

He showed* that the process of growth is intimately connected with the formation of ascending concentration currents feeding the growing crystal, so that faces on the lower side of the crystal grow faster than those above, and in order to secure regularity of growth he devised a form of apparatus in which the crystal was placed in the centre of a cylindrical vessel and the whole rotated slowly about a horizontal axis, so as to expose the several faces equally to the ascending currents.† This has formed the basis of the crystal-growing apparatus adopted.

For the measurement of growth, Wulff cut the crystal across at right angles to the faces to be measured, and stained the surface of the section so as to distinguish the layer formed from the nucleus on which it grew. This method limited the choice of material to episomorphous substances, and to those where differences of colour, either natural or producible by staining, distinctly mark the boundary between the nucleus and the crust formed. The crystal section was placed horizontally on the travelling stage of a microscope and the crust measured. In the present experiments this procedure was replaced by a method devised by Prof. H. L. Bowman, which had the double advantage of being applicable to all substances and of not destroying the crystal. The greater uniformity of the results obtained throughout the present experiments, as compared with those of Wulff, has been due as much to accuracy of measurement as to the special control of conditions affecting growth.

The form of apparatus adopted for growing the crystals was a modification of Wulff's rotating apparatus. The crystal was placed at the centre of a spherical flask filled with 500 c.c. of solution, the zone of faces to be measured being horizontal, and the flask was rotated slowly on a horizontal axis, parallel to that of the zone.

In preparing a crystal for an experiment a hole was bored into it with a fine drill, along the axis of the zone of the faces to be measured, and a long fine pin

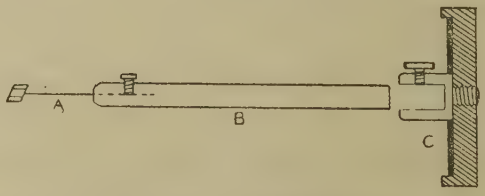


Fig. 8.—Crystal-holder.

was fixed into this with wax. This pin (A, fig. 8) was clamped by means of a set screw into one end of a cylindrical brass rod B, which served as a holder,

* 'Z. f. Kryst.,' vol. 34, p. 452 (1901).

† *Ibid.*, pp. 503, 508.

and the other end of which was fixed into a rubber-lined vulcanite cap C. This again was fitted on to the flask (D, fig. 9) so that the crystal was situated at

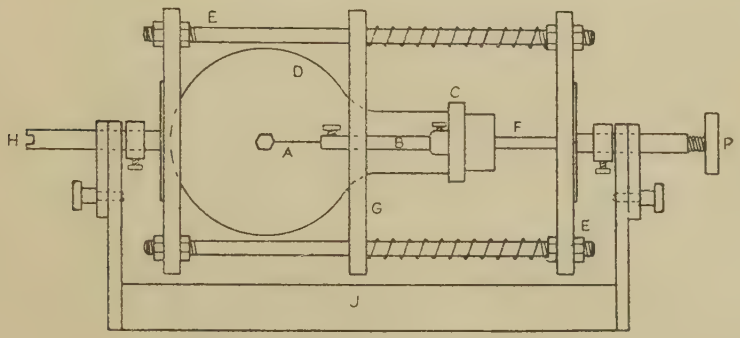


FIG. 9.—Apparatus for crystallisation.

the centre of the solution. The flask was mounted in a frame (E) by means of a wooden sliding piece (G) with springs, the cap being held firmly in position by the screw FP, and the whole was slowly rotated about the horizontal axis HP, supported on the stand J, by an electric motor with suitable reducing gear, connected to the clutch at H.*

During the rotation of the crystal about the horizontal axis every face parallel to that axis assumes successively all possible inclinations with reference to the ascending concentration currents, and this prevents the one-sided action of these currents on that zone. The mode of mounting the crystal, on a fine pin at the centre of a spherical flask, was intended to secure the least possible disturbance of these currents during the rotation.

For the purpose of measurement, after each period of growth the crystal-holder was removed from the vulcanite cap and placed on the two V-blocks (K) of a small tripod carrier (L, fig. 10), which rested on a slab of plate glass (Q) clamped to the stage of a microscope and had been previously adjusted by means of the levelling screw (M) so that the axis of the holder was parallel to, and at a fixed distance above the glass plate, on which the tripod was lightly pressed by the spring (N). The holder was held in the V's with slight friction by a second spring (O), and could be rotated by a large vulcanite milled head (P).

The zone having been adjusted parallel to the axis of the holder in the manner

* In Wulff's experiments the crystal was held in a 3-pointed brass clip which was fastened to the cover of a cylindrical crystallising vessel with sealing wax. The vessel containing the growing crystal was mounted on a frame and attached to the minute hand of a clock, so that it made one revolution about its horizontal axis in an hour.

described below, any face belonging to it could be set horizontal (*i.e.*, at right angles to the axis of the microscope) by means of the milled head, this position

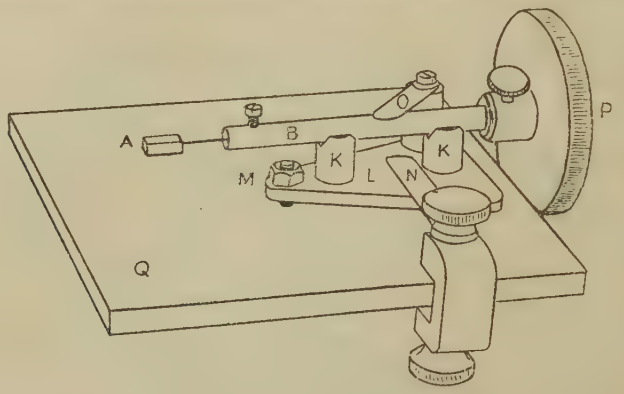


Fig. 10.—Measuring tripod.

being indicated by the fact that all parts of the face were in equally sharp focus when brought to the centre of the field by sliding the carrier on the glass plate. The magnification used was about 75 diameters.

Each face in turn was set horizontal and sharply focussed by means of the fine adjustment screw of the microscope (the coarse adjustment being left untouched and care being taken to avoid backlash), and the position of the graduated head (reading to 0·01 mm.) was noted. The crystal-holder was then removed from the carrier and replaced in the flask, filled with fresh solution, for a further period of growth, after which the readings for the several faces were repeated and the growth of each face (in millimetres) obtained by difference.

In this way a series of experiments could be made with the same crystal, each experiment giving independent results.

The proper adjustment of the crystal on the holder for a series of experiments was easily effected, provided reasonable care had been taken both in drilling the crystal and in mounting it on the pin. A face was focussed as described above and brought into the horizontal position partly by rotating the holder and partly by bending the pin by hand. A second face, as nearly as possible at right angles to the first, was similarly treated, the first face was corrected if necessary and after one or two trials a perfect adjustment of the zone was usually attained.*

* A slow tilting motion of the crystal (in one plane) can be obtained with the holder shown in fig. 11, in which the holder is drilled at a slight angle to the axis, and a notch is filed in the metal opposite the screw. The crystal can then be tilted by bending the pin (of hard brass wire) by means of the screw. In practice, however, the simple method described above was found to be sufficient.

A series of trial experiments was conducted to determine the most suitable conditions for regular growth, particularly as regards the concentration of the

Fig. 11.

solution and the rate of rotation of the flask. The actual crystal growth was always effected by the slow cooling of the solution. The initial temperature of saturation of the solution should not be too high on account of the difficulty in restraining the rate of cooling, and not too low in order that there may be a measurable amount of growth in a short time. The most suitable temperature had to be determined by actual experiments in every case, those finally adopted being as follows :—

Sodium chlorate	35 to 37° C.
Potash alum and double sulphates	30 to 34° C.
Potassium and ammonium tartrates.....	40° C.

The solution was put into the flask some 2° above its saturation point, to allow for initial cooling, and the crystal was thus introduced into a solution, which was just about saturated. This precaution was necessary in order to guard against any initial dissolution of the crystal. On the other hand, the introduction of a crystal into an already supersaturated solution is to be avoided, as it certainly leads to irregular growth.

The amount of solution used for each experiment was about 500 c.c., an amount sufficient to allow for the even growth of all forms. With small volumes of solution growth was restricted—a result which was noted by Weyberg.*

The flask was wrapped round with cotton wool to retard cooling, and the fall in temperature during an experiment lasting 4 to 5 hours was 10° to 12° C. The effect was tried of reducing still further the rate of cooling by immersing the apparatus in a tank of water initially heated to the temperature of the solution. The time of an experiment was thereby increased to 12 hours and the fall in temperature reduced to 5° or 6° C. The results obtained were good, but as satisfactory results were obtained from experiments of shorter duration and at a faster rate of cooling, this method was not continued.

The time allowed for an experiment depended somewhat on the velocity of growth of faces under observation, and care was taken that the series of experiments conducted with a single crystal were of equal duration. A period of

* 'Z. f. Kryst.,' vol. 34, p. 531 (1901).

growth of 4 to 5 hours was found to be most convenient; the amount deposited on a crystal being more than proportionately less during a shorter experiment, while a longer experiment gave excessive growth on a fast growing form, which might even become obliterated. The figures in the following table show the varying thicknesses of substance deposited on the faces of $p\{110\}$ and $c\{001\}$ of a crystal of magnesium-potassium sulphate in a consecutive series of experiments, and illustrate the effect of the duration of the experiment on the rate of growth observed.

Table I.—Effect of Duration of Experiment on the Rate of Growth.

Magnesium-potassium Sulphate.

No. of experiment.	Growth in millimetres.				Duration of experiment.
	$p(110)$	$c(001)$	$p(\bar{1}\bar{1}0)$	$c(00\bar{1})$	
					Hours
1	0·04	0·02	0·03	0·02	$1\frac{1}{2}$
2	0·05	0·04	0·04	0·04	$2\frac{1}{4}$
3	0·12	0·10	0·12	0·10	3
4	0·29	0·24	0·29	0·25	4

From the slow rate of growth observed during the first two experiments it might seem that the solution was undersaturated when the crystal was put in, and that the crystal had failed to grow during the first fall of temperature on that account. But even on close observation no rounded faces or bevelled edges could be seen. As a matter of fact a surprisingly small amount of growth was always recorded for similar experiments of short duration. Now Miers has shown that there are two conditions of supersaturation, the metastable and the labile. From a solution passing through the former condition, material is slowly deposited on crystals present. With a further fall of temperature the labile condition is reached; growth becomes more rapid, and in a stirred solution is accompanied by a shower of fine crystals. It would seem that the solution during the first stages in the above experiments passed through the metastable condition, and that the labile condition was reached after a fall in temperature of about 6°, the crystals then growing faster.

As the conditions controlling growth were never identical even for two successive experiments, the results of different experiments are not strictly comparable, but each experiment gave comparable results for the different faces.

The most satisfactory rate of rotation of the flask was about 4 or 5 turns per hour. If this rate was considerably increased there was too much movement in the solution, showers of small crystals tended to form and the original crystals grew irregularly. If the flask rotated much more slowly the results were less uniform.

3. *Determination of the Growth of a Crystal and Mode of Statement of the Results.*

The method employed is subject to the limitation that only one zone can be measured at a time, so that separate sets of experiments have to be made for the different zones, in turn. A proper choice of zones ultimately leads to a knowledge of the relative rates of growth, $K_{(hkl)} : K_{(pqr)}$, of all the forms present on the crystal.

At the beginning and end of an experiment readings for each face in the zone were taken with the microscope, as described on p. 66, and the difference between two successive readings for a given face gave the amount of growth in millimetres, as given in the tables. The first measurement was always taken after the crystal had been grown for some two or three hours under conditions similar to those of an ordinary experiment, to ensure a perfect smooth surface on all faces.

The number of experiments made for one zone varied, but (except in the preliminary experiments) was seldom less than 15. After each experiment the mean growth of the faces of each form was taken and the relative rates of growth of the different forms computed ; the mean of the results for each pair of forms being taken as the final value of their relative rate of growth, as in Table IV.

For reasons of space only a few such tables can be given here ; and in general the mean values only are stated, together with the range of variation in the different experiments, as an indication of the degree of accuracy attained.

Except where mentioned (Tables I, II, III, V) the results of all the experiments relating to a given zone are combined, whether made on the same or different crystals.

The results for the crystal as a whole are collected, for each substance, as in the Table on p. 73, in which the relative rates of growth of the several forms are expressed in terms of that of $c\{001\}$, taken as unity.

The uniformity of the results compares favourably with those of Wulff and Weyberg. The measured amounts of growth obtained varied in different cases between 0·05 mm. and 1·2 mm., but were commonly between 0·3 mm. and 0·8 mm. Those obtained in one experiment, on different faces of the same

form were usually equal within 0·01 or 0·02 mm., and (apart from the exceptional conditions described on p. 74) the difference between extreme values never exceeded 0·04 mm. In the values obtained in a series of experiments for the relative rate of growth of a given pair of forms, the maximum deviation from the mean seldom exceeded $\pm$ 5 per cent. and reached 10 per cent. in one case only.

4. *Preliminary Experiments with Alum and Sodium Chlorate.*

Experiments were first made with octahedral crystals of potash alum and with cubes of sodium chlorate, on which all the faces are of one kind, in order to test the reliability and accuracy of the methods employed, and to ascertain if the relative rate of growth of the faces was affected by the initial habit of the crystal.

On examining the results for a crystal of potash alum given in Table II, which are typical of many others, it will be seen that the rate of growth of different octahedral faces is the same. It was observed that an initial irregularity, however marked, in the habit of the crystal had no influence on the subsequent rate of growth of its faces. Thus, some crystals were originally of regular habit, while others were flattened parallel to a pair of faces; but the equal growth of all faces clearly demonstrated that habit does not affect the rate of growth when no other form is present.* Similar concordant results were obtained with cubic crystals of sodium chlorate. Here again only one form was present and all the faces grew equally, irrespective of any initial irregularity of habit.

A necessary consequence of equal growth is that an irregular or misshapen crystal *tends* to become regular, but only in the sense that the initial disparities become smaller fractions of the total dimensions. This is illustrated by the following results obtained with a tabular crystal of potash alum. At the beginning of the experiments its dimensions were about 6 mm. diameter $\times$ 2 mm. thick, and after having grown for 48 hours (a series of 12 experiments of 4 hours each) its size had increased to 12 mm. diameter $\times$ 8 mm. An equal thickness of material amounting to about 3 mm. had thus been deposited on each face, but the proportions of the crystal dimensions changed from 3 : 1 to 3 : 2. On other crystals similar initial proportions have changed to 4 : 3 and even to 5 : 4. Similar results were obtained with crystals of sodium chlorate.

* Compare p. 73.

Table II.—Potash Alum, $KAl(SO_4)_2 . 12H_2O$.

No. of experiment.	Growth in millimetres.			
	$o(111)$	$o(\bar{1}\bar{1}1)$	$o(\bar{1}\bar{1}\bar{1})$	$o(11\bar{1})$
1	0·43	0·42	0·43	0·43
2	0·63	0·61	0·61	0·63
3	0·61	0·61	0·63	0·62
4	0·43	0·44	0·45	0·43
Total growth	2·10	2·08	2·12	2·11

5. *The Double Sulphates.*

In the cubic substances above described the faces belonged to the same form. But with crystals of the monoclinic double sulphates numerous forms are present, not wholly included in any one zone, and the relative growth-rates of seven forms were determined, namely, $c\{001\}$, $b\{010\}$, $a\{100\}$, $p\{110\}$, $q\{011\}$, $r\{\bar{2}01\}$ and $o'\{\bar{1}11\}$. Although the total number of possible pairs of these forms is 21, the number of zones required to be measured was less than this, (1) because some zones contain more than two forms, and (2) because it was not necessary to measure directly all possible pairs.

The following are the zones actually measured :—

1. $p(110)$, $b(010)$, $a(100)$.
2. $p(110)$, $c(001)$, $o'(\bar{1}11)$.
3. $c(001)$, $q(011)$, $b(010)$.
4. $c(001)$, $r(\bar{2}01)$, $a(100)$.
5. $p(110)$, $r(\bar{2}01)$.
6. $p(110)$, $q(011)$.

The relative velocities of growth of all the forms could be deduced from the first four zones, and the final results were computed from the measurements made on them. Zones 5 and 6 were used for purposes of control. The measurements on these latter zones always confirmed the accuracy of the previous results.

The results obtained for crystals of magnesium-ammonium, iron-ammonium and magnesium-potassium sulphates will now be considered in detail. In each salt it will be convenient to describe the zones in the order mentioned in the preceding paragraph.

(*a*) *Magnesium-ammonium sulphate*, $Mg(NH_4)_2(SO_4)_2 . 6H_2O$.

Zone $p(110)$–$b(010)$.—The crystals showed considerable variation in habit, but many were examined which contained only the form $p\{110\}$ in this zone. The four faces of this form were found to grow at the same rate. Table III is typical and gives results for a series of nine successive experiments made with a single crystal. Fig. 12 illustrates the corresponding gradual change in habit, as seen in successive traces of the prism zone, and shows how the irregularity of a misshapen crystal becomes less marked during growth.

Table III.—Magnesium-ammonium sulphate, $Mg(NH_4)_2(SO_4)_2 . 6H_2O$.
Zone $[pp]$.

No. of experiment.	Growth in millimetres.			
	p (110)	p ($\bar{1}10$)	p ($\bar{1}\bar{1}0$)	p ($1\bar{1}0$)
1	0·34	0·35	0·35	0·35
2	0·42	0·44	0·44	0·42
3	0·44	0·42	0·43	0·44
4	0·28	0·29	0·29	0·28
5	0·25	0·26	0·25	0·26
6	0·22	0·23	0·22	0·22
7	0·45	0·45	0·46	0·45
8	0·47	0·46	0·46	0·46
9	0·51	0·51	0·50	0·52
Total growth	3·38	3·41	3·40	3·40

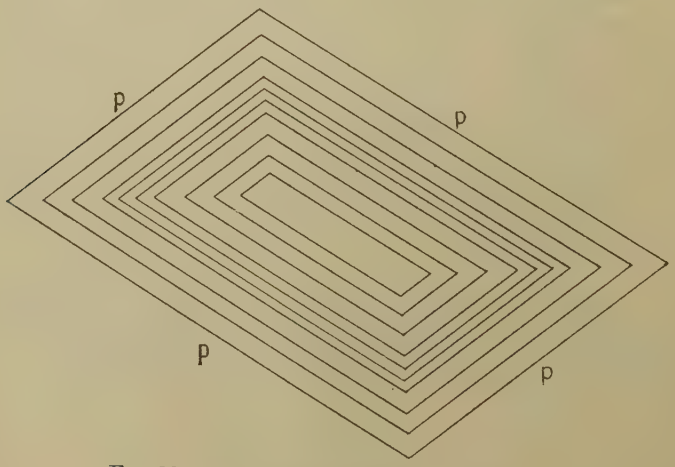


Fig. 12.—Magnesium-ammonium sulphate.

Certain crystals exhibited also a large face $b\{010\}$, and the ratio of growth $K_b : K_p$ was found to be 1·97 : 1,—b being, in fact, the fastest growing form observed.*

Zone $p(110)$–$c(001)$–$o'(\bar{1}\bar{1}1)$.—This zone was a very convenient one to measure, as p and c were always well developed; the face o' was small and sometimes absent. The results recorded in Table IV show that c grows faster than p in the ratio of 1 : 0·85, and slower than o' in the ratio of 1 : 1·12. After a series of experiments, c and p remained prominent forms, whilst o' persisted as a very small face or sometimes disappeared.

Table IV.—Magnesium-ammonium sulphate, $Mg(NH_4)_2(SO_4)_2 . 6H_2O$.
Zone $[p\ c\ o']$.

No. of experiment.	Growth in millimetres.						Relative rate of growth.	
	$c(00\bar{1})$	$p(\bar{1}\bar{1}0)$	$o'(\bar{1}\bar{1}1)$	$c(001)$	$p(110)$	$o'(11\bar{1})$	$K_p : K_c$.	$K_{o'} : K_c$.
1	0·06	0·05	—	0·06	0·05	—	0·83	—
2	0·41	0·33	—	0·41	0·34	—	0·82	—
3	0·27	0·23	—	0·28	0·24	—	0·85	—
4	0·45	0·37	0·49	0·45	0·38	0·50	0·83	1·10
5	0·44	0·36	0·48	0·44	0·36	0·48	0·82	1·09
6	0·49	0·42	—	0·50	0·42	—	0·85	—
7	0·52	0·45	0·60	0·52	0·46	0·59	0·87	1·14
8	0·61	0·50	0·69	0·61	0·53	0·71	0·84	1·15
9	0·50	0·43	0·56	0·50	0·44	0·56	0·87	1·12
10	0·44	0·38	0·52	0·46	0·40	0·50	0·87	1·13
11	0·53	0·45	—	0·54	0·47	—	0·86	—
12	0·31	0·28	—	0·31	0·26	—	0·87	—
13	0·27	0·24	—	0·30	0·24	—	0·84	—
14	0·22	0·18	0·25	0·23	0·19	0·25	0·82	1·11
15	0·27	0·24	0·31	0·27	0·23	0·29	0·87	1·11
16	0·38	0·32	0·43	0·37	0·31	0·41	0·84	1·12
17	0·49	0·41	—	0·48	0·41	—	0·85	—
18	0·17	0·15	—	0·17	0·13	—	0·82	—
						Mean =	0·85	1·12

Zone $c(001)$–$q(011)$–$b(010)$.—The measurements show that q grows a little faster than c in the ratio 1·10 : 1; and during growth this face persists as a prominent form. (The ratios obtained in 14 experiments varied between 1·03 and 1·16.)

In some crystals a large b-face was present and measurements show a mean growth-ratio $K_b : K_c = 1·66$ (varying in five experiments between 1·56 and

* Results obtained from certain excessively misshapen crystals of this substance are discussed below (p. 74).

1·71). The relative rate of growth of these faces can also be deduced from measurements made on other zones : for as $K_p : K_c = 0·85$, and $K_b : K_p = 1·97$, it follows that $K_b : K_c = 1·67$; and it will be seen that the direct results agree with the former measurements.

Zone $c(001)$–$r(\bar{2}01)$.—The results for this zone show that r grows slower than c in the ratio 0·86 : 1 (varying in 18 experiments between 0·82 and 0·90). After a series of experiments the r-face—which was always small—showed only a slight increase in size and did not become prominent. This is attributable to the fact that the adjacent faces of p have an equally slow rate of growth.

Control Measurements.—The above results were checked by control experiments carried out on the zones $[p\,r]$ and $[p\,q]$, and in each case the results accorded with the values previously obtained.

Thus, the ratio $K_p : K_r$, as calculated from values already given for $K_c : K_r$ and $K_c : K_p$, is 0·99, and this is in good agreement with a mean value 1·02 obtained from direct measurements (four experiments, 1·00 to 1·05).

Again, the growth-rate $K_q : K_p$, as computed for values already given for $K_c : K_p$ and $K_c : K_q$ is 1·29, which agrees (by chance exactly) with the mean obtained from direct measurements (four experiments, 1·22 to 1·34).

Final Results.—The following table for magnesium-ammonium sulphate gives the relative rates of growth of all the forms observed, that for $c\{001\}$ being taken as unity :—

Magnesium-ammonium sulphate, $Mg(NH_4)_2(SO_4)_2 . 6H_2O$.

Form.	$p\{110\}$	$r\{\bar{2}01\}$	$c\{001\}$	$q\{011\}$	$o'\{\bar{1}11\}$	$b\{010\}$
Relative rate of growth	0·85	0·86	1·00	1·10	1·12	1·67

Exceptional Behaviour of Certain Crystals.—Some crystals of the magnesium-ammonium salt were examined which had one prominent b-face and two prominent p-faces, as shown in fig. 1. The abnormal habit of such crystals was due to their having rested on the b-face during growth. The behaviour (which is typical) of one such crystal, which had a large face $b(0\bar{1}0)$, is shown in Table V, which gives the growth, during a series of seven experiments, of the faces in the prism zone. It will be seen from this and from the corresponding figure (fig. 13) that during the earlier experiments of the series the four prism faces do not grow equally. The two large prism faces (110) ($\bar{1}10$), remote from b, grow equally, and faster than the two small ones adjacent to b, which also show equal growth. The face b is a fast growing form, having a rate of growth nearly twice

Table V.—Magnesium-ammonium sulphate, $Mg(NH_4)_2(SO_4)_2 . 6H_2O$.
(Misshapen crystal.)

Zone [*p p b*].

No. of experiment.	Growth in millimetres.					Relative rate of growth.	
	p (110)	*p* ($\bar{1}10$)	*p* ($\bar{1}\bar{1}0$),	*b* ($0\bar{1}0$)	*p* ($1\bar{1}0$)	$K_{(0\bar{1}0)} : K_{(110)\,(\bar{1}10)}$	$K_{(1\bar{1}0)\,(\bar{1}\bar{1}0)} : K_{(110)\,(\bar{1}10)}$
1	0·33	0·32	0·25	0·65	0·26	2·00	0·78
2	0·17	0·17	0·14	0·32	0·13	1·88	0·80
3	0·12	0·13	0·10	0·24	0·11	1·92	0·84
						$K_b : K_p$	
4	0·20	0·20	0·19	0·38	0·20	1·93	0·98
5	0·15	0·15	0·14	0·28	0·14	1·93	0·93
6	0·29	0·29	0·29	0·58	0·31	1·97	1·03
7	0·20	0·20	0·21	0·40	0·20	1·98	1·02
Total growth	1·46	1·46	1·32	2·85	1·35		

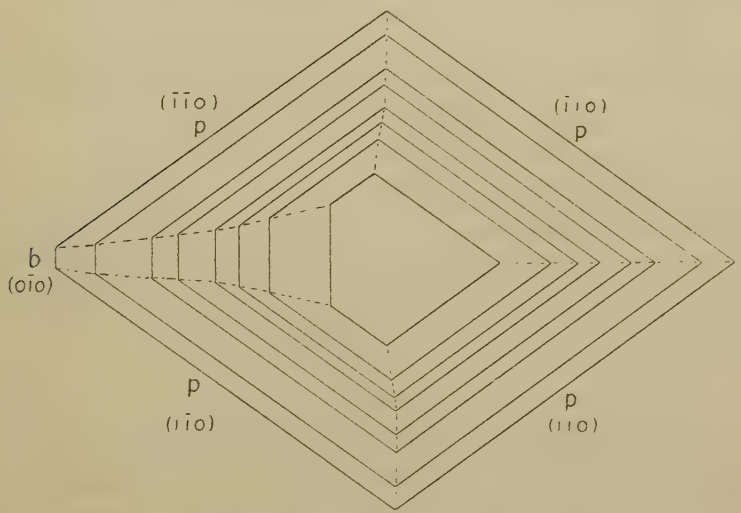


FIG. 13.—Magnesium-ammonium sulphate.

that of the remoter *p*-faces. In successive experiments, the difference between the rates of growth of the two sets of *p*-faces gradually diminishes, until they grow equally, but the *b*-face continues to grow at a uniform rate as compared with the *p*-faces remote from it. In the course of the experiments the crystal habit becomes modified, the *b*-face being quickly reduced in size until the crystal approaches the habit of a simple prism, as shown by the outlines of fig. 13.

The constancy, or otherwise, of the relative rates of growth of adjacent faces is well shown in the figure by the dotted lines passing through their intersections. For the direction of these lines changes with any change in the relative rates of growth.

The results obtained from this and similar crystals contrast so strongly with those previously given throughout this paper as to demand a brief examination. They establish the principle that a large face of an exceptionally fast growing form may influence the rate of growth of adjacent slower growing faces, drawing to itself material from the solution which under ordinary circumstances would be deposited on them. The precise limits within which this action occurs have not yet been thoroughly explored and much remains to be done in this and other directions. Thus although in the case quoted the rate of growth of the *b*-face $(0\bar{1}0)$ as compared with that of the remoter prism faces (110) $(\bar{1}10)$ did not markedly change, and the growth of the latter faces did not therefore appear to be affected by *b*, it would be unwise to rule out the possibility of such an effect in all cases.

The results so far obtained may thus be summarised :—

(1) A misshapen crystal completely bounded by faces of a single form (cube or octahedron) shows no tendency under the conditions of the experiments to correct its irregularity by unequal rates of growth on different faces, though with equal rates the irregularity becomes less marked (*cf.* fig. 12).

(2) The same is generally true of crystals bounded by different forms ; but (3) under very exceptional conditions an abnormally developed face having a high growth-rate may adversely influence its neighbours during subsequent growth under normal conditions, but only in the first stages, until its area is diminished by the extension of the adjacent faces.

(*b*) *Iron-ammonium sulphate*, $Fe(NH_4)_2(SO_4)_2 . 6H_2O$.

Zone *p*(110)–*b*(010).—The prism zone in this salt, which sometimes showed the two *b*-faces as well as the primary prism, was always weakly developed. Fifteen experiments on crystals having *b*-faces gave the mean relative rate of growth $K_b : K_p$ as 2·70. Owing to its high rate of growth, the *b*-face often becomes cloudy and its growth is somewhat irregular.

Zone *p*(110)–*c*(001)–*o*′$(\bar{1}\bar{1}1)$.—The results of 15 experiments show that *p* grows slower than *c* in the ratio 0·86 : 1, while in 10 experiments, where *o*′ was present, this form was found to grow faster than *c* in the ratio of 1·20 : 1. Throughout a series of experiments *c* persisted as a prominent form, while *o*′

remained small. (The observed values of $K_p : K_c$ were between 0·82 and 0·88; those of $K_{o'} : K_c$ were between 1·17 and 1·25.)

Zone $c(001)$–$q(011)$–$b(010)$.—The mean of 17 experiments showed that q grows considerably faster than c, in the ratio 1·24 : 1* (the observed values lying between 1·20 and 1·27). The form q is prominent in small crystals, and during growth it remains well marked though becoming relatively smaller. Three experiments where b was present gave a mean ratio $K_b : K_c = 2·37$ (2·34 to 2·40).

The rate of growth of b, measured in this zone, also acts as a check on the previous results, for the directly observed value of $K_b : K_c$ (2·37) agrees well with the value 2·32 computed from those obtained for $K_b : K_p$ and $K_p : K_c$.

Zone $c(001)$–$r(\bar{2}01)$.—In this salt $r\{\bar{2}01\}$ was generally the most prominent form, although it was occasionally small. Crystals were measured with r-faces of varying size, but the size did not affect the uniform slow growth of this form. This was the slowest growing form observed; its mean rate of growth being only 0·54 that of c. (In 22 experiments the values observed lay between 0·50 and 0·58.) During an experiment, r remained the largest form on a crystal, or (if small at first) quickly became prominent.

Control Measurements.—The results obtained in the above zones were confirmed by control experiments carried out on the zones $[p\,r]$ and $[p\,q]$.

Thus the ratio $K_r : K_p$ as deduced from values already given for $K_c : K_p$ and $K_c : K_r$ is 0·63, which is in satisfactory agreement with a mean value 0·59 obtained from direct measurements on this zone (three experiments giving 0·59, 0·58, 0·59).

Again, the ratio $K_q : K_p$ as computed from values already given for $K_c : K_p$ and $K_c : K_q$, is 1·44, which is identical with the mean result of direct measurements (five experiments giving values between 1·41 and 1·47).

Final Results.—The following table for Mohr's salt gives the relative rates of growth of all the forms observed, that of $c\{001\}$ being taken as unity :—

Iron-ammonium sulphate, $Fe(NH_4)_2(SO_4)_2 . 6H_2O$.

Form.	$r\{\bar{2}01\}$	$p\{110\}$	$c\{001\}$	$o'\{\bar{1}11\}$	$q\{011\}$	$b\{010\}$
Relative rate of growth	0·54	0·86	1·00	1·20	1·24	2·32

* Wulff gave 1·23 and 1·20 as the mean values of this ratio obtained in two different sets of experiments.

(*c*) *Magnesium-potassium sulphate*, $MgK_2(SO_4)_2 . 6H_2O$.

Zone $p(110)$–$a(100)$.—The prism zone of this salt often exhibited the faces of *a*, whereas *b*-faces were not observed.

The results of 21 experiments on this zone show that the *p*-faces grow equally, and faster than *a* in the ratio 1 : 0·70 (the values observed lying between 0·67 and 0·77). At first the *a*-faces were narrow, truncating edges of the prism, but after a series of experiments, in which their rate of growth was uniform, they became as large as the prism faces.

Zone $p(110)$–$c(001)$–$o'(\bar{1}11)$.—The results of 26 experiments on this zone show that *p* grows faster than *c* in the ratio 1·19 : 1 (the values obtained lying between 1·15 and 1·23). The face *o'* was not observed but the *c*- and *p*-faces were prominent.

Zone $c(001)$-$q(011)$–$b(010)$.—The relative rates of growth of *q* and *c* were found from 15 experiments to be as 1·36 : 1, from which it will be seen that *q* is a fast growing form. (The observed values varied between 1·33 and 1·40.) It is present as a small face and sometimes disappears with growth.

Zone $c(001)$–$a(100)$–$r(\bar{2}01)$.—The *r*-face grows faster than *c* in the ratio 1·66 : 1. When present this face is always small and tends to disappear with growth. The mean ratio $K_a : K_c$ was found to be 0·82. (In 11 experiments the observed values of $K_r : K_c$ varied between 1·60 and 1·74 ; those for $K_a : K_c$ between 0·78 and 0·86.)

Incidentally, the results obtained from this zone also act as a check on the rate of growth of *a*, for the directly observed value of $K_a : K_c$ (0·82) agrees well with that computed from the zones $[p\,a]$ and $[p\,c]$ (viz., 0·83).

Final Results.—The following table gives the relative rates of growth of all the forms observed in magnesium-potassium sulphate, that of *c*{001} being taken as unity :—

Magnesium-potassium sulphate, $MgK_2(SO_4)_2 . 6H_2O$.

Form.	*a*{100}	*c*{001}	*p*{110}	*q*{011}	*r*{$\bar{2}$01}
Relative rate of growth	0·83	1·00	1·19	1·36	1·66

(*d*) *Comparison of the Double Sulphates.*

The relative rates of growth of the forms observed in each of the three double sulphates, that of *c*{001} being taken as unity, are collected in Table VI to facilitate a comparison.

Table VI.

Form.	Magnesium-ammonium sulphate.	Iron-ammonium sulphate.	Magnesium-potassium sulphate.
$c\{001\}$	1·00	1·00	1·00
$a\{100\}$	—	—	0·83
$b\{010\}$	1·67	2·32	—
$p\{110\}$	0·85	0·86	1·19
$q\{011\}$	1·10	1·24	1·36
$r\{\bar{2}01\}$	0·86	0·54	1·66
$o'\{\bar{1}11\}$	1·12	1·20	—

From this general table it will be seen that for any given form, say $q\{011\}$, there is throughout the series no constancy in the rate of growth as compared with that of $c\{001\}$. Moreover, the order of increasing rates of growth of different forms is markedly different in the three salts. Thus, r, for example, is in Mohr's salt the slowest, in magnesium-potassium sulphate the fastest growing form.

A comparative study of each zone in the three double sulphates was made with special reference to changes in habit. The relative rates of growth of forms determines the habit of a crystal, for it determines the ultimate position occupied by any face on a growing crystal. It is always the slow growing forms which tend to persist and become prominent while rapidly growing faces gradually disappear, owing to the extension of surrounding faces.

In the prism zone the b-face was always the fastest growing form observed (its growth-rate reached its highest value in Mohr's salt). The a-face (only observed on magnesium-potassium sulphate) on the contrary grew slowly and assumed greater and greater prominence. The two figs. 14 and 15, showing

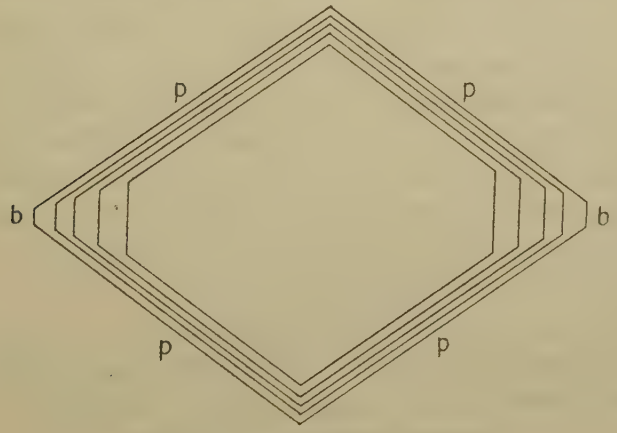


Fig. 14.—Iron-ammonium sulphate.

successive traces of the prism-zone during a series of experiments, illustrate the change in habit produced by the growth of the *b*- and *a*-faces in crystals

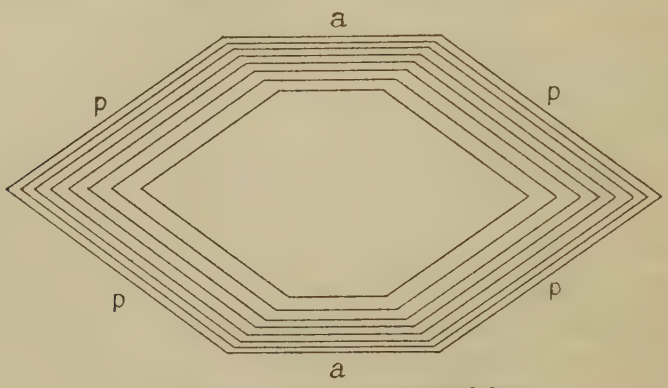


Fig. 15.—Magnesium-potassium sulphate.

of iron-ammonium sulphate and magnesium-potassium sulphate respectively. In both crystals the initial habit becomes modified, in the one case by the rapid decrease of *b*, leading to a simple prismatic habit, in the other by the increase of *a*, modifying the original prism. In fig. 13 can be seen the change in habit in the prism zone of a crystal of magnesium-ammonium sulphate. The large *b*-face tends to diminish by the extension of the prisms and the habit gradually becomes more regular.

In such a zone as [*p c o'*] on the other hand, growth caused little change in habit; the *p*- and *c*-faces are large and relatively slow growing, and persist as prominent forms. The face *o'* is fast growing and tends to disappear, but as it is always very small it does not appreciably affect the crystal habit. In magnesium-ammonium sulphate, in which the rate of growth of *p* is less than that of *c*, the crystal tends to become prismatic in habit, terminated by *c*. Where, as in the magnesium-potassium salt, *p* grows faster than *c* the crystal tends to become tabular parallel to *c*, the *p*-faces becoming relatively smaller.

In the zone [*c q b*] the size of *q* is closely related to its rate of growth. It is most prominent in the magnesium-ammonium salt, where its rate of growth is slowest ($K_q : K_c = 1 \cdot 10$) and the two forms *c* and *q* are about equally developed. In magnesium-potassium sulphate, where its rate of growth is considerably faster ($K_q : K_c = 1 \cdot 36$) it is small and often absent.

In the zone [*c r*], the *r*-face shows the importance of rate of growth as governing the size and persistence of a form. In the magnesium-potassium salt, this face was often absent or at the most very small. Its rate of growth was considerably faster than that of *c* and during an experiment it often disappeared.

In the magnesium-ammonium salt the *r*-faces were always small and remained as a small form after a series of experiments, only showing a slight increase in size, its rate of growth being 0·86 that of *c* and about the same as that of the adjacent *p*-faces (0·85). In Mohr's salt on the other hand, it was actually the slowest growing form observed, its rate of growth being only 0·54 that of *c*, and figs. 16 and 17 illustrate how, with growth, *r* remains, or quickly becomes, the largest form on the crystal.

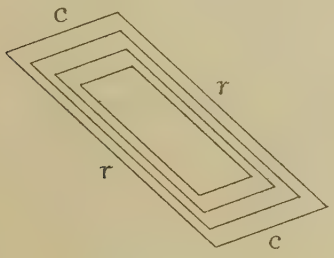


Fig. 16.

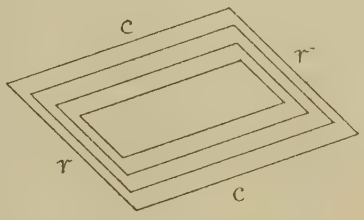


Fig. 17.

Iron-ammonium sulphate.

6. *The Neutral Tartrates.*

For the experiments on polar crystals, devoid of centrosymmetry, the monoclinic salts potassium tartrate, containing half a molecule of water of crystallisation, and ammonium tartrate (anhydrous) were used. Both salts crystallise well and exhibit a variety of forms terminating the *b*-axis. As the interest of these salts was centred in the relative rates of growth of the forms developed at opposite ends of the symmetry axis, measurements were limited to a single zone containing these forms.

In potassium tartrate (see fig. 6, p. 62) the zone included the large face $b'(0\bar{1}0)$ truncating one end, and the sphenoid $o(111)$ which formed the other end. The results for this zone appear in Table VII, from which it will be seen that the two faces of the sphenoid grow equally, and very much faster than $b'\{0\bar{1}0\}$, the ratio being 2·69 : 1.

In the absence of a face $b(010)$ the rate of growth normal to this face cannot be directly determined; but the elongation of the crystal in this direction due to the growth of the sphenoid $o\{111\}$ can be calculated from the growth-rate of this form, and it may be expected that the growth-rate of the *b*-face would be greater than this, so that if present it would tend to diminish and disappear.

Table VII.—Potassium tartrate, $K_2C_4H_4O_6 \cdot \frac{1}{2}H_2O$.

No. of experiment.	Growth in millimetres.			Relative rate of growth $K_o : K_{b'}$
	$b'(0\bar{1}0)$	$o(111)$	$o(\bar{1}1\bar{1})$	
1	0·18	0·47	0·45	2·56
2	0·19	0·51	0·51	2·68
3	0·13	0·34	0·35	2·65
4	0·22	0·59	0·59	2·68
5	0·20	0·56	0·53	2·72
6	0·06	0·15	0·17	2·67
7	0·09	0·26	0·24	2·78
8	0·06	0·15	0·16	2·58
9	0·08	0·21	0·21	2·63
10	0·09	0·25	0·26	2·83
11	0·17	0·46	0·48	2·77
12	0·05	0·13	0·14	2·70
13	0·06	0·16	0·16	2·67
14	0·09	0·23	0·25	2·67
15	0·17	0·42	0·45	2·56
16	0·06	0·16	0·17	2·75
17	0·09	0·24	0·24	2·67
18	0·11	0·28	0·31	2·68
19	0·04	0·11	0·11	2·75
20	0·14	0·38	0·40	2·79
21	0·08	0·23	0·21	2·75
22	0·18	0·47	0·49	2·67
				Mean = 2·69

From fig. 18 it will be seen that

$$ab = \frac{ao}{\cos bao} = \frac{a'b' \times 2{\cdot}69}{\cos 22° 35'} = a'b' \times 2{\cdot}91,$$

i.e., the elongation at the positive end of the *b*-axis is 2·91 times the growth of $b'(0\bar{1}0)$. Hence the growth-ratio $K_b : K_{b'}$ will be greater than 2·91.

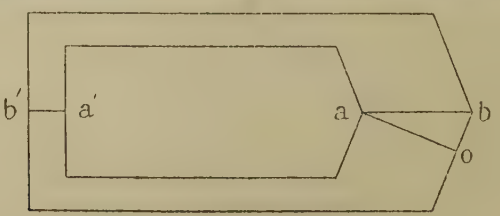


Fig. 18.—Potassium tartrate.

In view of the very different habit at the two ends of the crystal, it might be supposed that the result is influenced by asymmetry of habit (*cf.* p. 76). In order to avoid this source of error a face *b*(010) might have been produced on the crystal artificially by grinding, but it was unnecessary to resort to this

method as crystals of ammonium tartrate, although polar in character, exhibit a similar habit at the two ends (see fig. 7, p. 62). On this salt $q\{011\}$ and $q'\{0\bar{1}1\}$ were prominent forms. On some crystals other facets were also developed, sometimes differently at the two ends of the crystals; but as these were always quite small it cannot be supposed that their presence or absence had any appreciable effect on the rates of growth of q and q'. The results for the zone containing $c(001)$, $q(011)$ and $q'(0\bar{1}1)$ are given in Table VIII. The rate of growth of the two q-faces is the same, and differs from that of the two q'-faces, which also grow equally. The growth-ratios $K_q : K_c$ and $K_{q'} : K_c$ are 1·86 and 1·37 respectively, so that $K_q : K_{q'} = 1{\cdot}36$. Such a difference in the rates of growth of parallel faces situated at the two ends of the polar axis, the faces being of similar size and having the same inclination to the axis, can only be due to structural polarity.

Table VIII.—Ammonium tartrate, $(NH_4)_2C_4H_4O_6$.

No. of experiment.	Growth in millimetres.						Relative rate of growth.	
	$c(001)$	$q(011)$	$q(01\bar{1})$	$c(00\bar{1})$	$q'(0\bar{1}\bar{1})$	$q'(0\bar{1}1)$	$K_q : K_c$	$K_{q'} : K_c$
1	0·90	1·71	1·71	0·90	1·23	1·25	1·90	1·38
2	0·12	0·21	0·23	0·12	0·16	0·16	1·83	1·33
3	0·17	0·32	0·32	0·17	0·25	0·25	1·88	1·47
4	0·59	1·10	1·10	0·61	0·80	0·81	1·83	1·34
5	0·14	0·27	0·25	0·14	0·18	0·20	1·86	1·36
6	0·23	0·42	0·43	0·23	0·33	0·32	1·85	1·41
7	0·26	0·52	0·51	0·27	0·34	0·36	1·94	1·32
8	0·42	0·79	0·79	0·42	0·56	0·56	1·88	1·33
9	0·21	0·39	0·38	0·21	0·30	0·30	1·83	1·43
10	0·42	0·78	0·79	0·41	0·56	0·57	1·89	1·36
11	0·34	0·61	0·61	0·34	0·46	0·45	1·80	1·34
12	0·35	0·63	0·65	0·35	0·47	0·47	1·83	1·34
13	0·26	0·51	0·50	0·27	0·34	0·36	1·91	1·32
14	0·42	0·77	0·77	0·42	0·56	0·57	1·83	1·35
15	0·54	0·98	0·97	0·54	0·74	0·74	1·81	1·37
16	0·18	0·34	0·34	0·19	0·26	0·26	1·84	1·41
17	0·20	0·37	0·38	0·20	0·29	0·28	1·87	1·42
18	0·25	0·47	0·47	0·25	0·34	0·34	1·88	1·36
19	0·16	0·30	0·29	0·16	0·22	0·22	1·84	1·37
						Mean =	1·86	1·37

III.—Discussion of Results.

One of the main results of the present investigation is the proof that the various faces of a crystal grown under given uniform conditions have characteristic relative rates of growth. This fact has obviously an important bearing on

the habit of the crystal, and throws light (if only incompletely) on its past history and determines its future development under the given conditions.

The form of a crystal would seem to be determined by three main factors: (1) its structure, which is evidently determined by its chemical composition and the prevailing physical conditions of temperature and pressure; (2) the nature of the solution from which it separates; and (3) a variety of accidental circumstances, not easily subjected to exact study, such as its proximity to the walls of the enclosing vessel, or to other crystals, and the play of diffusion or concentration currents. The first two factors conjoined may be expected to lead to a standard or *ideal* habit to which individuals tend to conform in so far as they are not modified by the disturbing influences of the third category; and the immediate result of the present work is to justify this expectation and to establish the ideal type for each of the three double sulphates investigated.

This ideal habit is that which a crystal would assume if it could be subjected to uniform conditions of growth from the first moment of its formation. It can be obtained by constructing planes, parallel to the various faces observed on the actual crystal, at distances from a central point proportional to their measured rates of growth.

In such an ideal crystal those fast growing faces which, if present, would tend to diminish in size and disappear with growth will not be represented. The habit of the ideal crystal will not change with growth, but all the faces will extend proportionately, retaining their relative size and shape.

In practice, the ideal form has been constructed zone by zone, from the measurements, the ideal outline of each zone being found by drawing the traces of the several planes at distances from a centre proportional to their rates of growth. Any fast growing planes which tend to diminish by the extension of other faces in the zone will thus be eliminated. With this procedure it may happen that certain faces which form part of the ideal outline of one zone may in another zone be eliminated by the extension of adjacent faces of that zone. Such faces will not appear on the ideal crystal.

The ideal type corresponding to the conditions of the experiments is shown in figs. 19 to 21 for magnesium-ammonium, iron-ammonium and magnesium-potassium sulphates respectively.* A comparison with figs. 1 to 5 will illustrate in each case the difference between an actual crystal and the ideal type. It will be clear that practical considerations render it impossible to prepare an absolutely ideal crystal by growing it from the beginning under the

* It may be mentioned that a new point of view has been adopted in these three drawings, such that both o'-faces are at the back of the crystal in the upper half.

standard conditions, for there is a lower limit of size below which it cannot be handled, and an upper limit above which it becomes unmanageable.

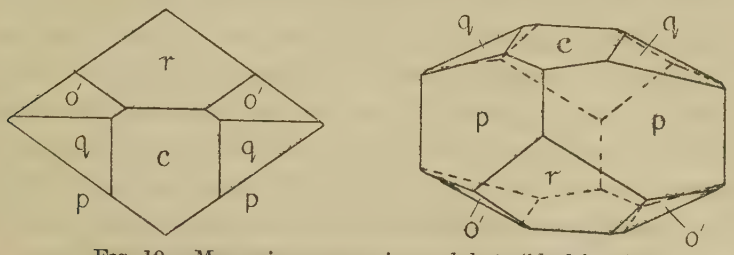


Fig. 19.—Magnesium-ammonium sulphate (ideal form).

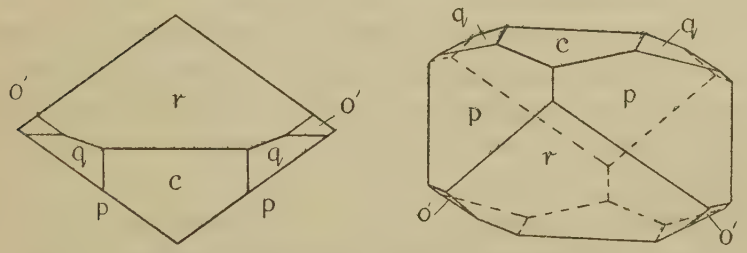


Fig. 20.—Iron-ammonium sulphate (ideal form).

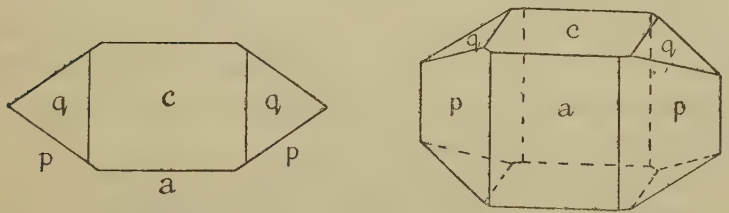


Fig. 21.—Magnesium-potassium sulphate (ideal form).

But it may be noted that an approach towards the ideal is more readily attainable in one respect than another. Rare faces, of fast growing forms such as $p'\{120\}$ and $b\{010\}$, easily outgrow themselves and vanish completely ; but an abnormal development in size of one or more of the common faces is not so quickly corrected by growth under uniform conditions, for it has been shown that relative rates of growth are constant, and it follows that the utmost a crystal can do is to approach, without quite attaining, the ideal habit (*cf.* p. 75). The rate at which such an approximation occurs is evidently determined by the relative velocities of growth of the various forms.

An apparent contradiction to the constancy of relative rates of growth was observed in the exceptional behaviour of certain crystals of magnesium-ammonium sulphate which originally exhibited a single dominant *b*-face and

in which the presence of this face influenced the rates of growth of the adjacent pair of prism faces. It must, however, be noted that this exceptional action of the *b*-face rapidly decreases with growth, as its area diminishes. When the adjacent *p*-faces become a little larger in proportion and the *b*-face smaller, the former acquire their normal relative rate of growth. The two large *p*-faces, remote from *b*, appear to grow at the normal rate throughout the whole experiment, and their rate maintains a constant ratio to that of *b*. It would appear, then, that in the early stages of growth the rapid deposit of material on the large *b*-face causes an impoverishment of the solution in its neighbourhood and draws in by diffusion or by concentration currents material which would normally be deposited on the adjacent faces.

While it is not proposed here to attempt a general discussion of the relations between rate of growth and crystal structure, attention may be drawn to two results of the experiments which seem to prove that the rate of growth of a face cannot be dependent in a simple maner on its reticular density as suggested by Wulff.

(1) It has been mentioned (p. 80) that in the isomorphous series of double sulphates examined the order of increasing rates of growth for different forms is widely different in the different salts, *r* being in one case the fastest and in another the slowest growing form observed. In such closely related salts, however, the space lattices must be strictly analogous in type and very similar in their dimensions, as indicated by the topic axes, so that the order of increasing reticular densities must be the same in all.

(2) More striking still is the marked difference in the rates of growth of parallel faces situated at opposite ends of the axis of symmetry in potassium and ammonium tartrates, since in such faces the reticular densities of the space lattice planes must be identical, and the difference must therefore be due to a polarity in the arrangement of the atoms about the points of the lattice.

Summary.

1. A method has been developed by which the relative rates of growth of the faces of a crystal may be measured with accuracy, the crystal being grown in a rotating vessel as suggested by G. Wulff.

2. By an extensive series of measurements on crystals of the isomorphous double sulphates of magnesium-ammonium, iron-ammonium and magnesium-potassium, of the formula $R''R'_2(SO_4)_2 \cdot 6H_2O$, of potassium tartrate, $K_2C_4H_4O_6 \cdot \frac{1}{2}H_2O$, and ammonium tartrate, $(NH_4)_2C_4H_4O_6$, the relative rates of growth of the faces occurring on these substances have been determined.

3. The measurements prove that under the conditions of the experiments described :—

(i) The similar faces of a *simple* form grow at the same rate, even when of different sizes. Hence a misshapen crystal, if grown under uniform conditions, tends towards, but never attains, the ideal form with equal faces.

(ii) On a combination :

Unlike faces grow at different rates.

Like faces grow at the same rate, except when adjacent to a large face of another, fast growing, form which causes impoverishment of the solution in its neighbourhood and destroys the uniformity of the conditions. Except in this case there is a constant ratio (under the given conditions) between the rates of growth of any two different forms.

(iii) In the case of crystals having no centre of symmetry the rates of growth of parallel faces may be widely different.

(iv) In an isomorphous series of salts the order of increasing rates of growth of different forms is not the same in different members of the series.

4. For each of the sulphates mentioned an "ideal form" has been constructed, which is that which a crystal would possess whose faces had grown from the beginning at the observed rates.

In conclusion I wish to acknowledge my indebtedness to Dr. T. V. Barker and Prof. H. L. Bowman, who both suggested the subject of the research and assisted me throughout with kind advice and criticism.

The work was carried out in the Mineralogical Department of the University Museum at Oxford; the cost of the equipment was defrayed by a grant from the Royal Society.

The Latent Heat of Vaporization of Benzene at Temperatures above the Boiling Point.

By J. A. Sutcliffe, F. C. Lay, and W. Ll. Prichard.

(Communicated by D. L. Chapman, F.R.S.—Received September 7, 1926.—Revised March 22, 1927.)

The progress made in developing the molecular theory of matter in the liquid state has been retarded by the circumstance that the experimental data, available for the purpose of testing theoretical conclusions, are scanty. For such a purpose the latent heat of the vapour of a suitable liquid would be particularly useful.

Numerous formulæ expressing the latent heat of vapours at high temperatures as a function of two or more of the physical constants of the vapour and liquid have been proposed. Some of the references to the best known of these are—

Bakker ('Z. f. Physik. Chem.,' vol. 18, p. 519, 1895); Crompton ('Proc. Chem. Soc.,' vol. 17, p. 61, 1901); Mills ('J. Phys. Chem.,' vol. 6, p. 209, 1902); Dieterici ('Ann. d. Physik,' vol. 35, p. 220, 1911; vol. 62, p. 75, 1920); Lewis ('Phil. Mag.,' vol. 25, p. 61, 1913); Appleby and Chapman ('J. Chem. Soc.,' p. 734, 1914); Partington ('Z. f. Physik. Chem.,' vol. 88, p. 29, 1914); Albertosi ('J. Chem. Phys.,' vol. 13, p. 379, 1915); Hammick ('Phil. Mag.,' vol. 41, p. 21, 1921); Thompson ('Chem. News,' vol. 123, p. 204, 1921); Mortimer ('J. Am. Chem. Soc.,' vol. 44, p. 1429, 1922).

A comparison of the latent heats deduced from such formulæ with experimental data is difficult because, with a few exceptions, such data are limited to the direct determinations of the latent heats at a single temperature, namely, the boiling point, and to values calculated by means of the Clausius-Clapeyron equation from S. Young's measurements of the vapour pressure and specific volume of the liquid and vapour of a number of organic liquids ('Proc. R. Dub. Soc.,' vol. 12, p. 374, 1910), *vide* Sutton ('Phil. Mag.,' vol. 29, p. 593, 1915). The degree of accuracy of the latent heat thus indirectly deduced is difficult to estimate and will not be known until the heat of vaporization of at least one of the liquids investigated by Young has been determined over a wide range of temperature. This communication is an account of a research conducted with this end in view.

Regnault (' Mém. de Paris,' 1862) determined the quantity of heat required to convert 1 gram of benzene liquid at 0° C. into saturated vapour at temperatures ranging from 7° C. to 215° C. The relation between this quantity of heat and the temperature of the saturated vapour was found to be in agreement with the empirical formula :—

$$\lambda = 109{\cdot}0 + 0{\cdot}24429t - 0{\cdot}0001315t^2,$$

in which t is the temperature and λ the specified quantity of heat. Unfortunately, data for the specific heat of benzene liquid are not available for the purpose of calculating the latent heat between these temperatures from Regnault's figures. It is clear, however, from the values obtained for the specific heat at 0° C. (Mills and MacRae, ' J. Phys. Chem.,' vol. 14, p. 797, 1910 ; Pickering ' Roy. Soc. Proc.,' vol. 49, p. 11, 1891) and the fact that this quantity increases with rising temperature, that the equation for the latent heat, L, would become :—

$$L = a - bt - ct^2 \ \ldots,$$

where a, b and c are all positive quantities. Thus the variation of the latent heat with temperature would be represented by a curve bending towards the temperature axis.

Griffiths and Marshall (' Phil. Mag.,' 1896) determined the latent heat of benzene at low pressures. The determinations were made over a comparatively small range of temperature, namely 30° C. to 50° C., and one additional experiment was performed at 20° C. Over this range the latent heat was found to be satisfactorily represented by the linear relation :—

$$L = 107{\cdot}05 - 0{\cdot}158t.$$

The value at the boiling point (80·2° C.) calculated from this formula is 94·37 calories per gram.

Direct experiment at the boiling point by several workers has furnished results which, taken as a whole, are in fair agreement. The following values have been obtained in recent years :—

Temperature.	Latent heat.	Experimenter.
80·0° C.	93·90	Nagornow & Rotinjanz (' Z. f. Physik. Chem.,' 1911).
80·2° C.	94·93	Brown (' J. Chem. Soc.,' 1905).
80·35° C.	94·35	Tyrer (' J. Chem. Soc.,' 1911).
80·2° C.	94·35	Matthews (' J. Am. Chem. Soc.,' 1926).

Matthews and Tyrer claim that their determinations at the boiling point are in agreement with the results of Griffiths and Marshall. The question

arises, however, as to whether the linear relation between the latent heat and the temperature, assumed by Griffiths and Marshall, will continue to hold at a temperature 30° C. above the highest temperature at which a determination was made. The experiments of Regnault show that the graph obtained by plotting the latent heat against the temperature is a curve bending towards the temperature axis. The authors have determined the latent heat of benzene vapour at temperatures between the boiling point and 152° C., and it will be shown in this communication that the curvature of the graph obtained by plotting the latent heat against the temperature cannot be neglected. Further, it will be shown that an empirical equation of the type

$$L = a - bt - ct^2 - dt^3,$$

where a, b, c and d are all positive quantities, furnishes values in agreement not only with the points on the curve experimentally determined at high pressures by the authors, but also with those obtained at low pressures by Griffiths and Marshall. Thus the value for the latent heat at the boiling point, deduced from the combined data of Griffiths and Marshall and of the authors, is appreciably lower than the value (94·37 calories per gram) calculated on the assumption that the linear relation between the latent heat and the temperature continues to hold accurately from 50° C. to the boiling point.

The experimental method was briefly as follows :—A measured quantity of electrical energy was supplied to a vacuum vessel (used as calorimeter) containing mercury, in which the bulb of a steel bomb was immersed. The benzene was allowed to evaporate from the bomb, through a fine steel capillary tube, at a rate so controlled, by manipulation of a needle valve at the top of the capillary tube, that the mean temperature of the calorimeter during the experiment was practically identical with that of a surrounding thermostat. The weight of benzene evaporated was determined by weighing the bomb before and after each experiment. The rate of evaporation was slow, about 8 gm. of benzene being evaporated per hour.

The design of the apparatus (constructed by the Oxford Scientific Instrument Co.) is shown in elevation in fig. 1. An iron tube (a) was supported inside a second tube (b) a short distance above the bottom. The space between the two tubes was filled with a suitable liquid. The "medicated paraffin," which was used for the bath when experiments were performed at temperatures well above the boiling point of benzene, became viscous at lower temperatures and was replaced by aniline. The liquid was maintained at a constant temperature by means of a thermo-regulator. A flat base was welded on to the outer tube.

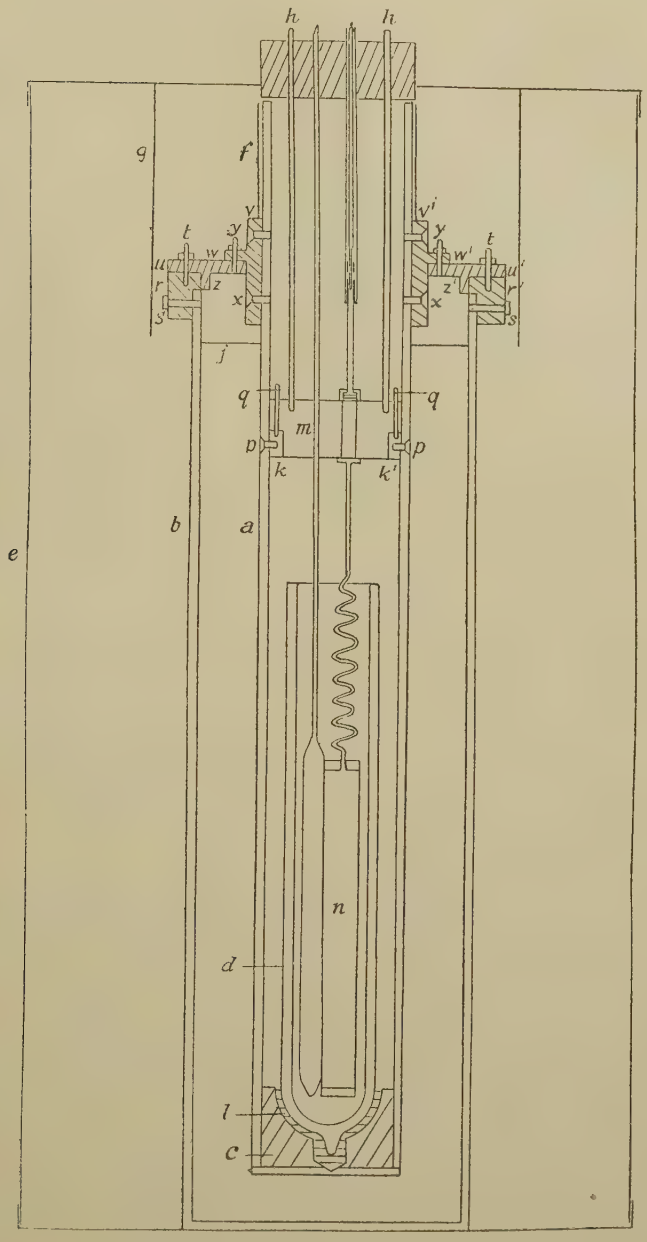


Fig. 1.

At the top of the tube a cast-iron ring (r, r') in which six pins (t) were screwed, was fixed firmly to the wall by screws (s). The vessel was covered by a cast-iron lid (u, u') bored with holes through which the pins passed. The lid was

secured to the cast-iron ring by nuts which were screwed down over the pins. In the lid four holes were drilled to take the top of the thermo-regulator, stirrer, heater and Beckmann thermometer, and in the centre a large hole was cut for the insertion of the inner tube.

The base (*c*), welded on to the end of this tube, was cut in a manner suitable for supporting the bottom of a vacuum vessel (*d*), used as calorimeter, which was wound with asbestos cord (*l*). At a convenient distance from the top of the tube a cast-iron collar (*v*, *v'*) with a flange (*w*, *w'*) was fixed to the wall by two parallel sets of screws (*x*). The flange was drilled to enable it to pass over six pins (*y*) screwed in the cast lid of the outer vessel and was held down by nuts.

The metal was cut from the flange at the places where it overlapped the holes in the cast lid. The whole apparatus was immersed in a cylindrical sheet-iron vessel (provided with a detachable lid, suitably bored with holes), which was packed with kieselguhr (*e*). To prevent the latter from entering the apparatus, steel tubes, which projected above the top of the kieselguhr, were tightly fitted into the holes in the cast lid.

Resistance wire was wound round the outer tube (*b*) and round two metal cylinders, one of which (*f*) fitted round that part of the inner tube which projected above the cast lid on the outer vessel, whilst the other (*g*) surrounded the upper portion of the apparatus. The heat generated by an electric current in the wire round the cylinder (*f*) served to keep the air space at the top of the inner tube at the correct temperature.

At a suitable distance down the inner tube (*a*), a copper ring (*k*, *k'*) was secured by screws (*p*). Resting on the ring was a block of copper (*m*) so cut that, when in position, the block and ring formed a complete cylinder. The ring contained three pins (*q*) over which the block passed and was thus prevented from rotating. The block was lowered on to the ring by two steel rods (*h*) screwed into the top and projecting above the lid of the vessel which contained the kieselguhr. The block was placed in a position well below the level of liquid (*j*) in the thermostat. Owing to the high thermal conductivity of copper, the block would acquire a uniform temperature, which would be nearly the same as that of the thermostat. In the block, holes were cut to admit the heater, thermometer, blowing tube and bomb. When these four objects were secured to it and immersed in the mercury contained in the calorimeter, the weight of the block more than counterbalanced the upthrust. The block served to prevent, as far as possible, variation in the conduction of heat to and from the calorimeter.

To construct the bomb (*n*) a steel rod was drilled longitudinally. A steel plug was screwed into each end of the tube thus formed and the edge of each plug was welded to the wall of the tube. The capacity of the bomb was about 64 c.c. The plug at the top of the bomb was drilled and tapped before it was welded and a steel capillary tube of 3 mm. external diameter, 0·5 mm. internal bore, and about 60 cm. in length, was screwed into it. The joint was then soldered with silver. The lower portion of the capillary tube was bent into a spiral so that when the bomb was placed in position in the calorimeter the lower portion of the spiral would be immersed in the mercury. The rest of the spiral served to increase the length of steel capillary tubing between the copper block and the mercury, whereby the heat conducted into the mercury was diminished.

The needle valve was 7 cm. above the top of the copper block. A square-cut piece of steel, soldered by silver to the capillary tube, fitted into a square hole in the block to which the bomb was secured. A glass tube was supported at a short distance below the needle valve. The benzene vapour, issuing from the valve, escaped from the apparatus through this tube. Inside the tube was a slender silver-steel screw key, secured to the needle valve, which could thus be manipulated outside the apparatus.

The bomb was electrolytically plated to prevent leakage. An unplated bomb was found to leak at the place where it was welded. Nickel was first used as coating, but it amalgamated with the mercury in the calorimeter at the high temperature. Hence the bomb was covered with electrolytic iron. A firm deposit of iron was obtained in the following way :—A strong solution of ferrous sulphate, boiled with magnesium powder for some time, was filtered through a Buchner funnel and used whilst hot as the electrolyte. During the process of electroplating, a small quantity of magnesium powder was introduced into the bath to prevent oxidation of the solution and the bomb was kept evacuated in order to remove gas from the capillary tubes through which the leak occurred.

Since the electrolytic coating rusted with great rapidity, the surface was covered with the adherent black magnetic oxide by the action of steam. After this treatment the weight alterations which occurred on immersing the bomb in the hot mercury were found to be negligible.

The needle valve was electrolytically coated with copper which, being a soft metal, effectually prevented any leak when the needle was pressed into the steel seating.

When the bomb was to be filled with benzene, a rubber stopper was pushed over the cylindrical container in order to support the bomb, with the capillary

tube pointing downwards, in a glass tube. The glass tube communicated by a condenser with a distillation vessel into which pure benzene, supplied by Kahlbaum, and some lumps of sodium were introduced. The nozzle of the bomb reached to the bottom of the glass tube in which it was supported. The desired quantity of dry benzene, which had been distilled over into this tube, was introduced into the bomb by evacuating the apparatus and then admitting air, dried with phosphorus pentoxide. Finally, the bomb was withdrawn from the apparatus and closed by the needle valve.

The cylindrical vacuum vessel, employed as calorimeter, was evacuated at a temperature of about 400° C. by Sprengel and mercury vapour pumps through a tube sealed on at the bottom of the outer wall. Unless evacuated at a high temperature, it rapidly became inefficient. When the objects were to be immersed in the calorimeter, the bomb was screwed to the copper block, the thermometer and heater were pushed through their respective holes in the block and bound to the bomb by wire. The block was then raised by the steel rods screwed into the top and carefully lowered into the apparatus. The rods were then unscrewed and removed. An asbestos lid covered the mouth of the calorimeter. A series of asbestos baffles were placed inside the inner tube to reduce the loss of heat from the air space at the top, and the mouth of the tube was covered by a box packed with asbestos wool.

The mercury in the calorimeter was stirred by a current of carbon dioxide blown into it through a glass capillary tube. The pressure of the carbon dioxide was kept constant by passing the gas from the cylinder into a large reservoir provided with a manometer on which the pressure could be read. The pressure required for the delivery of the desired quantity of gas to the calorimeter was considerably increased by inserting a length of very fine capillary tubing between the reservoir and the blowing tube. Thereby, small variations of pressure in the reservoir caused inappreciable alterations in the rate of flow of the gas into the calorimeter.

The heat generated by an electric current which passed through the wire wound round the outer iron tube (*b*, fig. 1) and the metal cylinder (*g*, fig. 1) maintained the bath at a temperature slightly below that desired. The small additional quantity of heat necessary to raise the bath to the required temperature was generated by an electric current passing through a resistance wire encased in a glass tube immersed in the liquid in the thermostat. A mercury-in-glass thermo-regulator which fitted into the annular space between the two tubes (*a* and *b*, fig. 1) operated in the usual manner by cutting off the current through the heater. The temperature of the bath was maintained

constant to about 0·01° C. The liquid was well stirred by a helical metal stirrer driven at high speed by an electric motor.

The air space at the top of the inner tube was maintained at a constant temperature slightly higher than that of the thermostat in order to prevent any condensation of benzene in the capillary tube of the bomb. To adjust the temperature to which this air space was heated, the following method was adopted:—The heaters of the air space, the top of the apparatus and the outer tube were connected in series with each other and with a rheostat, by means of which the total current, measured on a microammeter shunted off the circuit, was maintained constant. The relative currents through the heaters were suitably adjusted by rheostats shunted across them. A Beckmann thermometer, graduated in hundredths of a degree Centigrade, was used to indicate the temperature of the air space at the top of the inner tube.

For the construction of the thermometer used in the calorimeter, a large bulb, into which mercury was introduced, was sealed on to a long piece of very fine capillary tubing. A small bulb was fused to the top of the tube and sealed off. Precautions were taken to ensure that the mercury in the thermometer was free from gas bubbles. The mercury rose about 17 cm. up the capillary tube per degree Centigrade. Hence the thermometer was very sensitive. Nevertheless, the error caused by the sticking of the mercury to the walls of the glass capillary tube was not great, owing to the comparatively wide bore of the tube. In the outer bath a Beckmann thermometer, graduated in hundredths of a degree Centigrade, was used. This thermometer was calibrated by immersing it in the vapour of bromobenzene which was boiled under reduced pressure (in the case of the experiments performed at temperatures between 105° C. and 152° C.). The thermometer used in the calorimeter was immediately afterwards compared with the Beckmann thermometer in the thermostat. The vapour-pressure temperature curve for bromobenzene was determined using a thermometer standardized by the National Physical Laboratory, the pressures being read on a vacuum manometer. The temperature read from this curve for any value of the vapour pressure agreed to 0·1° C. with that quoted by Young ('J. Chem. Soc.,' vol. 47, p. 640, 1885). The temperatures given for the thermometer standardized by the National Physical Laboratory were on the constant volume air scale, accepting 444·5° C. as the boiling point of sulphur under normal pressure. At the lowest temperature at which experiments were performed water was used instead of bromobenzene, and the data of Regnault (Landolt-Börnstein tables, 1912 edition, p. 362) were accepted as correct. The operation of standardizing the thermometer used in the

calorimeter was performed both before and after each series of experiments. The results of the two standardizations agreed to 0·1° C., and the mean value was accepted as the correct temperature at which the series of experiments had been performed.

For the construction of the electric heater used in the calorimeter, a length of resistance wire, made of an alloy with a very small, negative, temperature coefficient, was wound in a spiral round a narrow glass tube. The two ends of the wire were fused into thick silver leads at the top of the tube. The leads were insulated by glass tubes through which they were pushed. The heater was encased in a wide glass tube into which paraffin was introduced. The level of the paraffin was several centimetres above the bottom of the silver leads but was slightly below the level of the mercury contained in the calorimeter.

The resistance of the heater was measured at 18° C. on a Wheatstone bridge by comparing it with standard coils supplied by W. G. Pye and Co., Cambridge. The temperature coefficient of the resistance wire, used for the construction of the heater, was also determined over the range of temperature at which experiments were performed. The diameter of the silver leads was measured and their resistance calculated from the following data of Dewar and Fleming ('Phil. Mag.,' vol. 36, p. 271, 1893; 'Roy. Inst. Gr. Brit.,' June 5, 1896).

Temperature.	Reciprocal resistance of a cubic centimetre.
0° C.	$68{\cdot}12 \times 10^4$
98·15° C.	$48{\cdot}49 \times 10^4$
192· 1° C.	$38{\cdot}34 \times 10^4$

The temperature coefficient of the resistance of the silver leads was also calculated from these data.

In order to apply a correction for the heat generated in the leads by the electric current, the assumption was made that all the heat generated below the mercury level entered the calorimeter, whilst all generated above the mercury level was lost.

The following values were thus obtained for the effective resistance of the heater :—

Temperature.	Resistance.
	Ohms.
86·86° C.	6·7176
105·42° C.	6·7124
125·25° C.	6·7069
143·49° C.	6·7017
151·80° C.	6·6994

The heater was standardized at the end of the first series of experiments and again at the end of the fourth series of experiments. The agreement between the results obtained showed that the resistance coil was free from secular changes.

The energy supplied to the calorimeter was calculated from the formula $H = C^2Rt/J$, where H is the heat in calories, C is the current in ampères, R is the resistance in ohms, t is the time in seconds, and J is the joule.

The current was measured by a silver voltameter. Two accumulators connected in series, each of 157 ampère-hours capacity, provided the electrical energy. A long platinum wire which dipped into mercury contained in a burette was inserted into the circuit. When mercury was added to the burette the length of platinum wire which was included in the circuit was altered. In this way the current, indicated on a galvanometer shunted off the circuit, was maintained constant. The following figures were taken for the calculations :— $J = 4·1895$ joules.* Electrochemical equivalent of silver $= 0·00111794$. The time was recorded by a stop-watch, which was checked repeatedly by means of the Greenwich time signal received by wireless telephony.

The method of conducting an experiment was as follows :—The calorimeter and contents were raised to a temperature approximately equal to that of the outer bath. When the temperature of the calorimeter was constant or altering at a very slow rate, the needle valve of the bomb was opened. The current through the electric heater in the calorimeter was switched on and the stop-watch was simultaneously started. The rate of evaporation of the benzene was controlled in such a manner that the mean temperature of the calorimeter was maintained, during the experiment, nearly the same as the original temperature. The rate of evaporation was slow. Hence the vapour escaped from the surface of the liquid without ebullition, and errors due to priming were

* This value was taken because the electrical methods of determining this constant have given higher values than that recommended by the International Critical Tables, and because it made possible a more exact comparison with the latent heat determinations of Griffiths and Marshall.

avoided. The temperatures of the calorimeter and the thermostat were recorded at intervals of two minutes throughout the period of the experiment (about three hours) so that a correction could be applied for radiation. At the end of the experiment the calorimeter was allowed to cool slightly below its initial temperature and the valve was closed. The current was switched off and the stop-watch simultaneously stopped when the temperature had risen to its former value. Usually the initial and final temperatures differed slightly, and a small correction was necessary. This correction, however, did not affect the latent heat determination by more than 0·1 per cent., even if the initial and final temperatures differed by as much as 0·01° C., which was seldom the case. The thermometer in the calorimeter was read until the temperature was constant or had assumed its original slow rate of alteration. The bomb was removed, cleaned and weighed.

It was necessary to apply a correction to the observed loss in weight of the bomb for two reasons :—

(1) A volume of benzene vapour equal to the volume of benzene liquid evaporated is left in the bomb.

(2) At the beginning of the experiment the bomb contains air, which is expelled by the prolonged evaporation.

Let W be observed loss in weight of bomb.
W_a be weight of air contained.
d_l be density of benzene liquid at the temperature of the experiment.
d_v be density of benzene vapour at the temperature of the experiment.
v be volume of liquid evaporated.
W_l be corrected weight of benzene evaporated.

Then

$$W - W_a = v\,(d_l - d_v).$$

Hence

$$W_l = vd_l = \frac{(W - W_a)\,d_l}{d_l - d_v}.$$

The values for d_l and d_v were obtained from Young's data (' Proc. Roy. Dub. Soc.,' 1910).

The method of counterpoise was adopted in performing the weighings. The weights were calibrated against a 1-gram weight standardized by Mr. J. J. Manley, M.A. The magnitude of the error introduced by not correcting the weights to *vacuo* was estimated as inappreciable for this particular case.

The rate of cooling of the calorimeter per minute per unit difference of

temperature between the calorimeter and the thermostat was determined by raising the temperature of the calorimeter and recording the rate of fall of the thermometer which it contained. The radiation correction was then calculated by the application of Simpson's rule. To determine the number of calories which must be added or subtracted (*a*) for radiation, (*b*) for the small rate of alteration (if any) of the temperature of the calorimeter before the experiment, and (*c*) for the difference between the initial and final temperatures of the calorimeter, the capacity for heat of the calorimeter is required. This was determined by an electrical method. The current was switched on through the heater for a measured time. A correction for radiation, etc., was applied to the resulting rise in temperature.

The following tables give the experimental results :—

First Series. Mean temperature = 151·80° C.

Number of experiment.	Latent heat.
(1)	79·747
(2)	79·951
(3)	80·025
(4)	79·938
(5)	79·798
(6)	79·726
(7)	79·818
	Mean 79·86 ± ·03

Second Series. Mean temperature = 143·49° C.

Number of experiment.	Latent heat.
(1)	81·751
(2)	81·766
(3)	81·728
(4)	81·700
	Mean 81·74 ± ·01

Third Series. Mean temperature = 125·25° C.

Number of experiment.	Latent heat.
(1)	85·501
(2)	85·513
(3)	85·609
(4)	85·492
	Mean 85·53 ± ·02

Fourth Series. Mean temperature = 105·42° C.

Number of experiment.	Latent heat.
(1)	89·185
(2)	89·228
(3)	89·281
(4)	89·111
	Mean 89·20 ±·02

Fifth Series. Mean temperature = 86·86° C.

Number of experiment.	Latent heat.
(1)	92·742
(2)	92·740
(3)	92·668
(4)	92·715
	Mean 92·72 ± ·01

These tables may be summarized as follows :—

Temperature.	Heat of vaporization.	Number of experiments performed.
86·86	92·72	4
105·42	89·20	4
125·25	85·53	4
143·49	81·74	4
151·80	79·86	7

The results obtained by Griffiths and Marshall may be summarized as follows :—

Temperature.	Heat of vaporization.	Number of experiments performed.
19·95	103·82	1
30·07	102·30	6
40·05	100·71	11
50·01	99·14	8

The most probable values of the four constants, a, b, c, and d in the equation

$$L = a + bt + ct^2 + dt^3,$$

in which L is the heat of vaporization in calories and t is the temperature,

have been calculated, using all the experimental results of Griffiths and Marshall and of the authors (49 experimental results).

The equation was found to be :—

$$L = 106{\cdot}868 - 1{\cdot}47106 \times 10^{-1}t - 1{\cdot}46582 \times 10^{-4}t^2 - 3{\cdot}60266 \times 10^{-7}t^3.$$

The values obtained from this equation and those obtained by direct experiment are shown in the following table. In the last column the values calculated by Mills (from Young's data on the vapour pressure and specific volumes of the liquid and vapour of benzene) by means of the Clausius-Clapeyron equation are given.

Temperature.	Heat of vaporization from the equation.	Heat of vaporization by experiment.	Experimenters.	Calculated value for heat of vaporization.
19·95	103·87	103·82	Griffiths & Marshall	—
30·07	102·30	102·30	,, ,, ...	—
40·05	100·72	100·71	,, ,, ...	—
50·01	99·10	99·14	,, ,, ...	—
86·86	92·75	92·72	Sutcliffe, Lay & Prichard	94·19
105·42	89·31	89·20	,, ,, ...	90·16
125·25	85·44	85·53	,, ,, ...	85·61
143·49	81·68	81·74	,, ,, ...	82·10
151·80	79·90	79·86	,, ,, ...	80·42

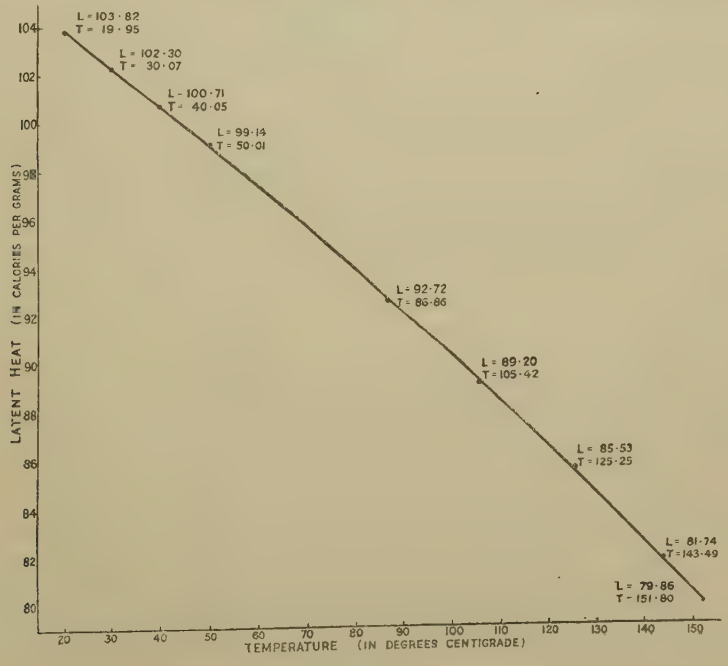


FIG. 2.

The value at the boiling point (80·2° C.) calculated from the equation is 93·94 calories per gram, in close agreement with the value obtained by Nagornow and Rotinjanz, but lower than that obtained by Tyrer and Matthews.

The graph obtained by plotting the latent heat against the temperature is shown in fig. 2.

Summary.

A method is described by which the heat of vaporization of liquids may be accurately determined at temperatures above their boiling points.

The method has been applied to the determination of the heat of vaporization of benzene. Consistent results have been obtained at five different temperatures.

These results are shown to be in agreement with the results of Griffiths and Marshall at temperatures below the boiling point. The values calculated from the equation :—

$L = 106{\cdot}868 - 1{\cdot}47106 \times 10^{-1}t - 1{\cdot}46582 \times 10^{-4}t^2 - 3{\cdot}60266 \times 10^{-7}t^3$,

in which L is the heat of vaporization and t is the temperature, are in close agreement with the results obtained by direct experiment at all the temperatures at which experiments have been performed by Griffiths and Marshall and by the authors.

In conclusion, one of us (J.A.S.) wishes to thank the Department of Scientific and Industrial Research for a grant which enabled him to continue the investigation for a further two years, during which period the results recorded above were obtained by him with the apparatus in its final form. The authors also desire to thank the Royal Society for a generous grant which defrayed part of the expenses of the apparatus, and Mr. D. L. Chapman for his advice and help during the progress of the investigation.

The Relations Connecting the Angle-Sums and Volume of a Polytope in Space of n Dimensions.

By D. M. Y. Sommerville, Victoria University College, Wellington, N.Z.

(Communicated by Major P. A. MacMahon, F.R.S.—Received November 29, 1926.)

§ 1. *Introduction.*

1.1. In two-dimensional spherical or elliptic geometry we have the familiar relation between the area of a triangle and its angle-sum,

$$\Delta = k^2(\Sigma\alpha - \pi), \tag{1.11}$$

2π being the measure of the whole angle at a point, and k the space-constant, or, in spherical geometry, the radius of the sphere. It is well known that in three-dimensional spherical or elliptic geometry there is no corresponding relation involving the volume of a tetrahedron.* For elliptic or hyperbolic space of four dimensions it was proved by Dehn† that the volume of a simplex can be expressed linearly in terms of the sums of the dihedral angles (angles at a face), angles at an edge, and angles at a vertex, but for space of five dimensions the linear relations do not involve the volume. He indicates also, in a general way, the extensions of these results for spaces of any odd or even dimensions. He shows further that these results are connected with the form of the Euler‡ polyhedral theorem, which is expressed by a linear relation connecting the numbers of boundaries of different dimensions, and which for space of odd dimensions is not homogeneous, *e.g.*, $N_2 - N_1 + N_0 = 2$ in three dimensions, but for space of even dimensions is homogeneous, *e.g.*, $N_1 - N_0 = 0$ in two dimensions, $N_3 - N_2 + N_1 - N_0 = 0$ in four. The connection, as Dehn points out, was made use of by Legendre in a proof which he gave for the Euler formula in three dimensions.§

* H. W. Richmond has investigated an expression for the volume of a tetrahedron in elliptic space in terms of an integral, 'Q. J. Math.,' vol. 34, p. 175 (1903). Lobachevsky himself investigated similarly the volume of a tetrahedron in hyperbolic space, "Foundations of geometry," 1829, and "Application of imaginary geometry to some integrals," 1836.

† M. Dehn, "Die Eulersche Formel im Zusammenhang mit dem Inhalt in der nicht-euklidischen Geometrie," 'Math. Ann.,' vol. 61, p. 561 (1905).

‡ L. Euler, 'Mém. Pétersb.,' 1758.

§ A. M. Legendre, "Eléments de géométrie," Liv. vii, Prop. xxv. (1794). The proof is reproduced in Todhunter's "Spherical Trigonometry," Chap. xii, or Todhunter and Leathem, Chap. xvi.

1.2. Dehn extends this connection in detail for four and five dimensions, and states the following general results in space of n dimensions R_n for simplexes and for polytopes bounded entirely by simplexes :

(1) In R_n there are $\frac{1}{2}n+1$ or $\frac{1}{2}(n+1)$ relations (according as n is even or odd) between the numbers of boundaries of a polytope bounded by simplexes.

(2) There is a linear relation between the various angle-sums of a simplex (a *Zerlegungsinvariante*) which does or does not involve the volume according as n is even or odd.

In the present paper these relations are actually obtained, and it is found for any polytope bounded by simplexes that the two kinds of relations, those which connect the numbers of boundaries, and those which connect the angle-sums, are of precisely the same form.

§ 2. *Measure of angles.*

2.1. We must first consider the different types of angles and their measure. In two dimensions there is just one type of angle to be considered, the angle between two directed straight lines or rays. Let O be the vertex; draw a circle with centre O, and let the two rays cut the circle in P, Q. The ratios of the arc PQ to the circumference, and the sector POQ to the area of the circle, are equal, and either may be taken as a measure of the angle. The whole angle at a point would then have the measure unity. In the *radian measure*, 2π is taken as the measure of the complete angle at a point; this is equivalent, in euclidean geometry, to taking as the measure of any angle the ratio of the arc to the radius.

2.2. In non-euclidean geometry the latter ratio is not constant for a given angle, but we may still take as the radian measure of the angle 2π times the ratio of the arc to the circumference, or 2π times the ratio of the sector to the area of the circle. In spherical geometry, if the radius is increased until the circle reduces to the point antipodal to O, the radian measure of the angle becomes 2π times the ratio of the whole area enclosed by the two rays to the whole area of the plane. In elliptic geometry, if the radius is increased until the circle becomes a straight line (the polar of O), the radian measure of the angle becomes 2π times the ratio of the whole area enclosed by the two rays to double the area of the plane. But if k is the space-constant the area of the whole plane is: in spherical geometry $4\pi k^2$, and in elliptic geometry $2\pi k^2$. In both spherical and elliptic geometry therefore the radian measure of an angle is the ratio of the area enclosed by the two rays to $2k^2$. (In spherical geometry the

rays begin at O and end at the antipodal point O′; in elliptic geometry they begin and end at O.)

2.3. In three dimensions it is usual to distinguish two types of angles: the angle between two planes (dihedral angle), and the angle between three or more planes (solid angle). These, however, are to be considered as of the same type. Let O be the vertex, or, in the case of the dihedral angle, any point on the edge, and draw a sphere with centre O. Then we may take as a measure of the angle the ratio of the area which the planes cut out on the surface of the sphere to the whole surface of the sphere, or the corresponding ratio of volumes. The measure of the whole angle at a point or an edge would then be unity. In euclidean geometry it is customary to take as the measure of a solid angle the ratio of the area cut out on the surface of the sphere to the square of the radius. We may call this the *radian measure*; the measure of the complete solid angle at a point would then be 4π. We may therefore define the radian measure of a solid or dihedral angle as 4π times the ratio of the area cut out by the bounding planes on the surface of the sphere to the whole surface of the sphere, or 4π times the corresponding ratio of volumes. In spherical geometry the ratio of the volume enclosed between the planes and the sphere to the whole volume of the sphere becomes in the extreme case the ratio of the whole volume enclosed between the planes (on a specified side of each of them) to the whole volume of space. As the whole volume of spherical space of space-constant k is equal to the hypersurface of a hypersphere of radius k, $= 2\pi^2 k^3$, we have: the radian measure of a solid or dihedral angle in spherical or elliptic geometry is the ratio of the volume enclosed between the planes to $\frac{1}{2}\pi k^3$.

2.4. Generally, in n dimensions there are angles bounded by $2, 3, \ldots, n-1$, or more than $n-1$, hyperplanes. In euclidean geometry we define the radian measure of an angle as the ratio of the hypersurface cut out of a hypersphere whose centre is on the axis (vertex, edge, etc.) to the $(n-1)$th power of the radius. The hypersurface of a hypersphere of radius k in n dimensions is $k^{n-1} n\pi^{\frac{1}{2}n} / \Gamma\{\frac{1}{2}(n+2)\}$. Hence the radian measure of the complete angle at a point is $n\pi^{\frac{1}{2}n} / \Gamma\{\frac{1}{2}(n+2)\}$. In spherical or elliptic geometry similarly the radian measure of an angle is the ratio of the hypervolume enclosed by the bounding hyperplanes to

$$\frac{k^n (n+1)\, \pi^{\frac{1}{2}(n+1)}}{\Gamma\{\frac{1}{2}(n+3)\}} \cdot \frac{\Gamma\{\frac{1}{2}(n+2)\}}{n\pi^{\frac{1}{2}n}}, \text{ i.e., to } k^n \frac{n+1}{n} \frac{\sqrt{\pi}\,\Gamma\{\frac{1}{2}(n+2)\}}{\Gamma\{\frac{1}{2}(n+3)\}}, \quad (2.41)$$

and the radian measure of the complete angle is

$$\frac{n\pi^{\frac{1}{2}n}}{\Gamma\{\frac{1}{2}(n+2)\}}. \quad (2.42)$$

The total volume of spherical space is

$$\frac{k^n (n+1)\, \pi^{\frac{1}{2}(n+1)}}{\Gamma\{\frac{1}{2}(n+3)\}}. \tag{2.43}$$

2.5. In hyperbolic space also we may lay down corresponding measures of angles, although the total volumes are not now available. Thus we may define the radian measure of an angle as proportional to the volume or the surface cut out of the surrounding hypersphere, the numerical factor being so adjusted that the measure of the complete angle has the same value as in elliptic or euclidean space.

§3. *The relations connecting the angle-sums of a simplex in spherical space of n dimensions.*

3.1. A simplex in space of n dimensions is bounded by $n + 1$ hyperplanes, and, generally, ${}_{n+1}C_{r+1}$ r-dimensional boundaries.

A single hyperplane divides space into two regions, which may be regarded as the positive and the negative side of the hyperplane, and denoted by $+$ and $-$. Two hyperplanes divide space into four regions, which may be distinguished by the signs $++, --$; $+-, -+$, and fall into two pairs of opposite or antipodal regions. Three hyperplanes give eight regions, and finally the $n + 1$ hyperplanes divide space into 2^{n+1} regions, which may be distinguished by the $n + 1$ signs $+$ and $-$ taken in a definite order. Each of these regions is the interior of a simplex, and we shall take the $(n + 1)$ $+$ signs as denoting the interior of the particular simplex with which we are dealing.

Every set of n hyperplanes determines a pair of antipodal points. We shall denote the $n + 1$ vertices of the simplex by $0, 1, 2, \ldots, n$, the antipodal points being denoted by $0', 1', 2', \ldots, n'$. The vertices of the 2^{n+1} simplexes are represented by the $n + 1$ digits $0, 1, 2, \ldots, n$ with or without accents. But we shall find it convenient to suppress the accented figures. Thus the antipodal regions (interiors of simplexes) $(0'12\ldots n)$ and $(01'2'\ldots n')$ will be denoted by $(12\ldots n)$ and (0) respectively, showing their relation to the given simplex as standing on an $(n - 1)$-dimensional boundary and a vertex respectively. $(012\ldots n)$ is the interior of the given simplex, and $(\ )$ is the interior of the antipodal simplex.

The number of regions of the type $(012\ldots r)$ is ${}_{n+1}C_{r+1}$, and there are the same number of antipodal regions $(r + 1, r + 2, \ldots, n)$. When n is even, antipodal regions are always of different type, but when n is odd the antipodal regions $\{01\ldots\frac{1}{2}(n - 1)\}$ and $\{\frac{1}{2}(n + 1), \ldots, n\}$, which both involve $\frac{1}{2}(n + 1)$ figures, are of the same type. Thus, in three dimensions, there are 4 regions

on the faces of a tetrahedron, and the antipodal regions are the 4 regions on the vertices, while the 6 regions on the edges form three pairs of antipodal regions. We shall denote any region of the same type as $(01\ldots r)$ by $[r+1]$.

3.2. We shall now consider the angular regions contained by any number of the bounding hyperplanes. The positive side of a hyperplane has already been implicitly defined as that region which contains the interior of the simplex, the region $(012\ldots n)$. The interior of any angular region is that which contains the interior of the simplex, or the region on the positive side of each of the bounding hyperplanes, and the antipodal region is that on the negative side of each of the bounding hyperplanes; we shall call the latter the *co-interior* of the angular space. We shall denote the content or volume of the interior of the angular space bounded by the two hyperplanes $123\ldots n$ and $023\ldots n$ by $\alpha_{23\ldots n}$, and that of the co-interior by $\alpha'_{23\ldots n}$; thus the number of suffixes is equal to the dimensions of the *axis* of the angular region. $\alpha_{23\ldots n}$ contains all those regions whose symbols include both 0 and 1, $\alpha'_{23\ldots n}$ all those whose symbols exclude both 0 and 1. Similarly $\alpha_{34\ldots n}$ denotes the content of the interior of the angular space bounded by the three hyperplanes $123\ldots n$, $023\ldots n$, $013\ldots n$, and contains all those regions whose symbols include 0, 1 and 2; and so on. $\alpha_{12\ldots n}$ will be taken to mean the whole region on the positive side of the hyperplane $12\ldots n$, $\alpha'_{12\ldots n}$ the negative side; *i.e.*, $\alpha_{12\,.\,n}$ is equal to the half of space $= \frac{1}{2}S = \alpha'_{12\ldots n}$. $\alpha_{01\ldots n}$ can be taken to mean the whole of space $= S$. α will be taken to denote the volume V of the interior of the simplex, α' that of the antipodal simplex.

3.3. The number of angular spaces of the type $\alpha_{r+1,\ldots n}$ is ${}_{n+1}C_{r+1} = {}_{n+1}C_{n-r}$. This angle contains the following regions:

1 region $(012\ldots n)$,
${}_{n-r}C_1$ regions whose symbols are formed with n of the digits
0, 1, ..., n always including 0, 1, 2, ..., r,
${}_{n-r}C_2$ regions with $n-1$ digits, and so on,
1 region $(012\ldots r)$.

3.4. We next consider the sums of the angular regions of the same type. Let $\Sigma[r]$ be denoted by A_r, so that A_{n+1} denotes the volume of the simplex V, and A_0 the (equal) volume of the antipodal simplex. Let A'_r denote the sum of the regions antipodal to those which compose A_r. Then $A'_{r+1} = A_{n-r}$, both in total volume and in separate parts, and $A_{r+1} = A_{n-r}$ in volume.

Let $\Sigma\alpha_{01\ldots r}$, the summation extending to all angular regions with $r+1$ suffixes, be denoted by Σ_r, $\Sigma\alpha'_{01\ldots r}$ by Σ'_r. $\Sigma\alpha\,(=\alpha)$ may be denoted by

Σ_{-1} so that $\Sigma_{-1} = V$. $\Sigma\alpha_{01\ldots n} = \alpha_{01\ldots n} = \Sigma_n$ = the whole volume of space S. $\Sigma\alpha_{12\ldots n} = \Sigma_{n-1} = \frac{1}{2}(n+1)\,S = \frac{1}{2}(n+1)\,\Sigma_n$.

Now a given region $[s]$ $(s \geqq n-r)$ is contained in each of the angular regions of type $\alpha_{01\ldots r}$ whose suffixes contain all the $n+1-s$ numbers which are not included in the symbol of $[s]$, and the number of these angular regions is ${}_sC_{n-r}$; hence in the sum $\Sigma\alpha_{01\ldots r}$ each region $[s]$ occurs ${}_sC_{n-r}$ times. Hence

$$\begin{aligned}\Sigma_r &= {}_{n+1}C_{n-r}\,A_{n+1} + {}_nC_{n-r}\,A_n + \ldots + {}_{n-r}C_{n-r}\,A_{n-r}\\ &= {}_{n+1}C_{r+1}\,V + {}_nC_r\,A_1 + \ldots + {}_{n-r}C_0\,A_{r+1}. \qquad (3.41)\end{aligned}$$

Thus, putting $r = 0, 1, \ldots, n$, we have the following $n+1$ equations in $A_1, \ldots, A_n$.

$$\begin{aligned}\Sigma_0 &= {}_{n+1}C_1\,V + A_1.\\ \Sigma_1 &= {}_{n+1}C_2\,V + {}_nC_1\,A_1 + A_2,\\ &\ldots\ldots\ldots\ldots\\ \Sigma_{n-1} &= {}_{n+1}C_n\,V + {}_nC_{n-1}\,A_1 + {}_{n-1}C_{n-2}\,A_2 + \ldots + A_n,\\ \Sigma_n &= V + A_1 + A_2 + \ldots + A_n + V.\end{aligned}$$

Also the relations

$$A_{r+1} = A_{n-r}\{r = 0, 1, \ldots, \tfrac{1}{2}n - 1 \text{ or } \tfrac{1}{2}(n-1)\} \qquad (3.42)$$

supply $\frac{1}{2}n$ or $\frac{1}{2}(n+1)$ further equations, according as n is even or odd. Hence by eliminating the A's we get, connecting the volume V and the angle-sums Σ_r, $\frac{1}{2}n + 1$ or $\frac{1}{2}(n+1)$ relations, *i.e.*, $[\frac{1}{2}n] + 1$, where $[\frac{1}{2}n]$ denotes as usual the integral part of $\frac{1}{2}n$.

3.5. We may form the eliminants by solving the equations (3.41) for the A's and substituting in (3.42).

From (3.41) we have

$$\begin{aligned}A_1 &= \Sigma_0 - {}_{n+1}C_1\,V,\\ A_2 &= \Sigma_1 - {}_nC_1\,\Sigma_0 + {}_{n+1}C_2\,V,\\ A_3 &= \Sigma_2 - {}_{n-1}C_1\,\Sigma_1 + {}_nC_2\,\Sigma_0 - {}_{n+1}C_3\,V.\end{aligned}$$

By induction we may show that

$$\begin{aligned}A_r = \Sigma_{r-1} - {}_{n-r+2}C_1\,\Sigma_{r-2} + \ldots + (-)^{r-s-1}\,{}_{n-s}C_{r-s-1}\,\Sigma_s + \ldots\\ + (-)^{r-1}\,{}_nC_{r-1}\,\Sigma_0 + (-)^r\,{}_{n+1}C_r\,V, \qquad (3.51)\end{aligned}$$

for assuming this true up to r, we have by (3.41)

$$A_{r+1} = \Sigma_r - {}_{n+1}C_{r+1}\,V - {}_nC_r\,A_1 - {}_{n-1}C_{r-1}\,A_2 - \ldots - {}_{n-r+1}C_1\,A_r.$$

Substituting the values of $A_1, \ldots, A_r$ in the right-hand side, the coefficient of Σ_s is

$$-{}_{n-s}C_{r-s} + {}_{n-s}C_{r-s} \cdot {}_{n-s}C_1 - \ldots + (-)^{p+1}\,{}_{n-s-p}C_{r-s-p} \cdot {}_{n-s}C_p + \ldots + (-)^{r-s}\,{}_{n-r+1}C_1 \cdot {}_{n-s}C_{r-s-1}.$$

But

$${}_{n-s-p}C_{r-s-p} \cdot {}_{n-s}C_p = {}_{n-s}C_{r-s} \cdot {}_{r-s}C_p.$$

Hence the coefficient of Σ_s is

$$-{}_{n-s}C_{r-s}\{1 - {}_{r-s}C_1 + \ldots + (-)^{r-s-1}\,{}_{r-s}C_{r-s-1}\} = (-)^{r-s}\,{}_{n-s}C_{r-s},$$

which establishes (3.51).

Hence, finally, equating A_r and A_{n-r+1} we have

$$\begin{aligned}\Sigma_{r-1} - {}_{n-r+2}C_1\,\Sigma_{r-2} + \ldots + (-)^{r-s-1}\,{}_{n-s}C_{r-s-1}\,\Sigma_s + \ldots \\ + (-)^{r-1}\,{}_nC_{r-1}\,\Sigma_0 + (-)^r\,{}_{n+1}C_r\,V \\ = \Sigma_{n-r} - {}_{r+1}C_1\,\Sigma_{n-r-1} + \ldots + (-)^{n-r-s}\,{}_{n-s}C_{n-r-s}\,\Sigma_s + \ldots \\ + (-)^{n-r}\,{}_nC_{n-r}\Sigma_0 + (-)^{n-r+1}\,{}_{n+1}C_{n-r+1}\,V, \quad (3.52)\end{aligned}$$

as the linear relations connecting the volume and angle-sums of a simplex. We notice that *if n is odd*, V *disappears from all the equations.*

The functions on the two sides of this equation occur frequently. We shall define the notation

$${}_n\phi_s(\Sigma) \equiv {}_{n-s}C_0\,\Sigma_{s-1} - {}_{n-s+1}C_1\,\Sigma_{s-2} + \ldots + (-)^s\,{}_nC_s\,\Sigma_{-1}. \quad (3.53)$$

Then the equations (3.52) can be written

$${}_{n+1}\phi_r(\Sigma) = {}_{n+1}\phi_{n-r+1}(\Sigma). \quad (3.54)$$

For $r = n + 1$ or 0,

$$\Sigma_n - \Sigma_{n-1} + \ldots + (-)^n\,\Sigma_0 + (-)^{n+1}\,V = V. \quad (3.55)$$

We have also the simple relation, already noted in 3.4,

$$\Sigma_{n-1} = \tfrac{1}{2}(n+1)\,\Sigma_n, \quad (3.56)$$

which can also be derived by elimination. In fact, since $A_1 = A_n$, $A_2 = A_{n-1}$, etc., the last two equations of (3.41) give

$$\begin{aligned}2\Sigma_{n-1} &= 2(n+1)\,V + (nA_1 + A_n) + \{(n-1)\,A_2 + 2A_{n-1}\} + \ldots \\ &= (n+1)\,(2V + A_1 + A_2 + \ldots) = (n+1)\,\Sigma_n.\end{aligned}$$

3.6. The foregoing investigation applies to spherical geometry. In elliptic geometry antipodal points coincide, the regions $(01\ldots r)$ and $(r+1, \ldots, n)$ are

continuously connected and form one region, and the sums A_{r+1} and A_{n-r} are coincident instead of being merely of equal volume. But instead of Σ_n we have $2\Sigma_n$. The final equations are therefore the same except for this modification.

3.7. Now let S_r denote the sum of all the angles at r-dimensional edges expressed in radian measure, so that

$$\Sigma_r = S_r \cdot k^n \frac{n+1}{n} \frac{\sqrt{\pi}\Gamma\{\frac{1}{2}(n+2)\}}{\Gamma\{\frac{1}{2}(n+3)\}}, \quad (r = 0, 1, \ldots, n-2) \tag{3.71}$$

and

$$\Sigma_n = k^n \frac{(n+1)\pi^{\frac{1}{2}(n+1)}}{\Gamma\{\frac{1}{2}(n+3)\}}, \tag{3.72}$$

$$\Sigma_{n-1} = \tfrac{1}{2}(n+1)\Sigma_n. \tag{3.73}$$

Equation (3.55) then becomes

$$\{1+(-1)^n\}V = k^n \frac{n+1}{n} \frac{\sqrt{\pi}\Gamma\{\frac{1}{2}(n+2)\}}{\Gamma\{\frac{1}{2}(n+3)\}} \left\{S_{n-2} - S_{n-3} + \ldots + (-)^n S_0 - \tfrac{1}{2}n(n-1)\frac{\pi^{\frac{1}{2}n}}{\Gamma\{\frac{1}{2}(n+2)\}}\right\},$$

and the other equations of (3.52) can be transformed similarly. The equations thus expressed in radian measure are the same in spherical and elliptic geometry.

Excluding the equation $\Sigma_{n-1} = \frac{1}{2}(n+1)\Sigma_n$ we have therefore $[\frac{1}{2}n]$ linear equations connecting the volume V and the $n-1$ angle-sums, and in these V disappears when n is odd.

The following are the equations up to $n = 4$:—

$$n = 2 \qquad V = k^2(S_0 - \pi),$$

$$n = 3 \qquad S_0 - S_1 + 4\pi = 0,$$

$$n = 4\ (*) \qquad V = \tfrac{2}{3}k^4(S_2 - S_1 + S_0 - 3\pi^2)$$

$$V = \tfrac{2}{15}k^4(2S_2 - 3S_1 + 5S_0 - 5\pi^2).$$

* From these equations we derive

$$-3V/k^4 = S_2 - 2S_0 - 4\pi^2,$$

and

$$0 = S_2 - \tfrac{2}{3}S_1 - \tfrac{10}{3}\pi^2.$$

The expressions on the right are, with different notation, Dehn's "Zerlegungsinvarianten" (*loc. cit.*, p. 572). With a disregard for homogeneity he takes different units for the angles at a face and the angles at a vertex or an edge, viz., for the complete dihedral angle at a face the value 2π and for the complete angle at a vertex or an edge the value unity. In our formulæ the complete angle in each case has the value $2\pi^2$, the surface content of a hypersphere of radius unity in euclidean space of four dimensions, *i.e.*, the total volume of spherical space of three dimensions and space-constant unity.

3.8. In euclidean geometry the space-constant $k \longrightarrow \infty$; the resulting equations are equivalent to putting $V = 0$, giving in all dimensions only linear identities involving the angle-sums. For hyperbolic geometry k is a pure imaginary ; when n is odd V and k disappear, and the relations between the angle-sums are the same in all three geometries ; when n is even k occurs to an even power ; *e.g.*, for $n = 2$ in hyperbolic geometry, putting ik instead of k, we have

$$V = k^2(\pi - S_0),$$

while for $n = 4$ the equations are the same as in elliptic geometry.

§ 4. *The relations connecting the numbers of boundaries of different dimensions of a polytope in space of n dimensions.*

4.1. Let the number of r-dimensional boundaries be denoted by N_r, and let the number of p-dimensional boundaries which are incident with (pass through or lie in) a particular q-dimensional boundary be denoted by N_{pq}. If we sum the numbers N_{pq} for all the q-dimensional boundaries the sum is equal to the number of p-dimensional boundaries each counted as often as there are q-dimensional boundaries incident with it, and this is the same as the number of q-dimensional boundaries each counted as often as there are p-dimensional boundaries incident with it. Hence

$$\Sigma N_{pq} = \Sigma N_{qp}. \tag{4.11}$$

For a particular p-dimensional boundary Euler's formula for p dimensions gives

$$N_{p-1,p} - N_{p-2,p} + \ldots + (-)^{p-1} N_{0,p} + (-1)^p = 1.$$

Summing for all the p-dimensional boundaries

$$\Sigma N_{p-1,p} - \Sigma N_{p-2,p} + \ldots + (-)^{p-1} \Sigma N_{0,p} + (-)^p N_p = N_p. \tag{4.12}$$

4.2. Take any vertex and draw a small hypersphere round it. This is cut by the boundaries at the vertex in a hyperspherical polytope with N_{r0} $(r-1)$-dimensional boundaries. Hence by Euler's formula

$$N_{n-1,0} - N_{n-2,0} + \ldots + (-)^{n-3} N_{20} + (-)^{n-2} N_{10} + (-1)^{n-1} = 1,$$

and summing for all the vertices

$$\Sigma N_{n-1,0} - \Sigma N_{n-2,0} + \ldots + (-)^{n-2} \Sigma N_{10} + (-)^{n-1} N_0 = N_0.$$

Take next an edge and any point O on the edge. Draw a hyperplane through O perpendicular to the edge, and a hypersphere of $n-1$ dimensions with centre O and lying in the hyperplane. This hypersphere is cut by the

boundaries at the edge in a hyperspherical polytope of $n-2$ dimensions with N_{r1} $(r-2)$-dimensional boundaries. Hence by Euler's formula

$$N_{n-1,1} - N_{n-2,1} + \ldots + (-)^{n-3} N_{21} + (-1)^{n-2} = 1.$$

Summing for all the edges

$$\Sigma N_{n-1,1} - \Sigma N_{n-2,1} + \ldots + (-)^{n-3} \Sigma N_{21} + (-)^{n-2} N_1 = N_1.$$

Proceeding similarly with the p-dimensional boundaries we have

$$\Sigma N_{n-1,p} - \Sigma N_{n-2,p} + \ldots + (-)^{n-p-2} \Sigma N_{p+1,p} + (-)^{n-p-1} N_p = N_p. \quad (4.21)$$

These relations are not all independent, and when n is even it can be shown that Euler's formula for the whole polytope is derived from them algebraically; for taking the term in N_p to the right, multiplying all the equations (4.12) by $+1$ and equations (4.21) by -1 and adding, we have on the right

$$2(N_{n-1} - N_{n-2} + \ldots + N_1 - N_0),$$

and on the left we have the terms

$$(-)^{p-q-1} \Sigma N_{qp} - (-)^{n-p-q-1} \Sigma N_{pq} = 0.$$

Thus *Euler's theorem for even values of n follows from its truth for all smaller values of n and the equations* (4.11).

For odd values of n the equations do not involve the numbers N_{2r} at all as these disappear both from equations (4.12) and from equations (4.21). This is noted by Dehn for $n = 5$, and he establishes Euler's formula for $n = 5$ by dividing the polytope into simplexes.

4.3. *Polytope bounded entirely by simplexes.*—When the boundaries of the polytope are all simplexes we can derive relations connecting the numbers N_r alone, without the individual numbers N_{pq}. In this case we have

$$N_{pq} = {}_{q+1}C_{p+1} \qquad (p < q) \quad (4.31)$$

and

$$\Sigma N_{pq} = \Sigma N_{qp} = {}_{q+1}C_{p+1} N_q \qquad (p < q). \quad (4.32)$$

Equations (4.12) become identities, while equations (4.21) become

$${}_nC_{p+1} N_{n-1} - {}_{n-1}C_{p+1} N_{n-2} + \ldots + (-)^{n-p-2} {}_{p+2}C_{p+1} N_{p+1} + (-)^{n-p-1} N_p = N_p \qquad (p = 0, 1, \ldots, n-1), \quad (4.33)$$

For $p = n-1$, however, we get merely the identity $N_{n-1} = N_{n-1}$, and the equations for $p = n-2$ and $p = n-3$ both reduce to

$$nN_{n-1} = 2N_{n-2}. \quad (4.34)$$

If we assume also Euler's formula for the whole polytope, it may be considered as included in (4.33) for $p = -1$, with the understanding that $N_{-1} = 1$, thus

$$N_{n-1} - N_{n-2} + \ldots + (-)^{n-1} N_0 + (-1)^n = 1. \tag{4.35}$$

4.4. The equations (4.33) are in fact not all independent. They may be expressed in another form, which is sometimes more useful, by first eliminating N_{n-1} between the first two (Euler's equation (4.35) being counted as the first one), then N_{n-1} and N_{n-2} between the first three, and so on. Thus, multiplying the first two equations respectively by $-n$ and 1, and adding, we get

$$N_{n-2} - 2N_{n-3} + 3N_{n-4} - \ldots + (-)^{n-2}(n-1)N_0 + (-)^{n-1} n = N_0 - n.$$

Multiplying the first three equations by ${}_nC_2$, $-{}_{n-1}C_1$, and 1 and adding, we get

$$N_{n-3} - {}_3C_2 N_{n-4} + {}_4C_2 N_{n-5} - \ldots + (-)^{n-1} {}_{n-1}C_2 N_0 + (-)^n {}_nC_2 = N_1 - {}_{n-1}C_1 N_0 + {}_nC_2;$$

and, generally, multiplying the first $r+1$ equations by ${}_nC_{n-r}$, $-{}_{n-1}C_{n-r}$, ${}_{n-2}C_{n-r}$, ..., $(-1)^r$ respectively, and adding, we get

$$N_{n-r-1} - {}_{r+1}C_1 N_{n-r-2} + \ldots + (-)^{n-r-1} {}_{n-1}C_{n-r-1} N_0 + (-)^{n-r} {}_nC_{n-r} = N_{r-1} - {}_{n-r+1}C_1 N_{r-2} + \ldots + (-)^{r-1} {}_{n-1}C_{r-1} N_0 + (-)^r {}_nC_r, \tag{4.41}$$

that is,

$$ {}_n\phi_r(N) = {}_n\phi_{n-r}(N). \tag{4.42}$$

Thus *the equations which connect the number of boundaries of a polytope bounded entirely by simplexes are precisely the same as those which connect the volume and angular regions of a simplex, but in one dimension less.* The number of these equations is $[\frac{1}{2}(n+1)]$.

4.5. *Relations between the number of boundaries of a pyramid.*—We may interpolate here the special relations for a pyramid.

Consider a pyramid whose base is a polytope of $(n-1)$ dimensions with N'_r r-dimensional boundaries, and let N_r be the number of r-dimensional boundaries of the pyramid.

Then

$$N_0 = N'_0 + 1, \ldots, N_r = N'_r + N'_{r-1}, \ldots, N_{n-1} = 1 + N'_{n-2}.$$

Therefore

$$\begin{aligned}
N'_0 &= N_0 - 1, \\
N'_1 &= N_1 - N'_0 = N_1 - N_0 + 1, \\
&\ldots\ldots\ldots\ldots\ldots \\
N'_r &= N_r - N_{r-1} + N_{r-2} - \ldots + (-)^r N_0 + (-1)^{r+1}, \\
&\ldots\ldots\ldots\ldots\ldots \\
N'_{n-2} &= N_{n-2} - N_{n-3} + \ldots + (-)^{n-2} N_0 + (-1)^{n-1}, \\
1 &= N_{n-1} - N_{n-2} + \ldots + (-)^{n-1} N_0 + (-1)^n.
\end{aligned} \tag{4.51}$$

The last relation is the first Euler relation for the pyramid, viz., ${}_n\phi_n(\mathrm{N}) = 1$. This direct proof is of interest as it holds for odd or even dimensions and does not involve induction from lower dimensions or other extraneous assumptions.

We have also for the base of the pyramid

$$\begin{aligned} 1 &= \mathrm{N}'_{n-2} - \mathrm{N}'_{n-3} + \ldots + (-)^{n-2}\mathrm{N}'_0 + (-1)^{n-1} \\ &= \mathrm{N}_{n-2} - 2\mathrm{N}_{n-3} + \ldots + (-)^{n-2}(n-1)\,\mathrm{N}_0 + (-)^{n-1}\,n, \end{aligned}$$

i.e.,

$$ {}_n\phi_{n-1}(\mathrm{N}) = 1 \tag{4.52}$$

If the base is itself a pyramid of $(n-1)$ dimensions it follows that ${}_{n-1}\phi_{n-2}(\mathrm{N}') = 1$,

i.e., $\mathrm{N}'_{n-3} - 2\mathrm{N}'_{n-4} + \ldots + (-)^{n-3}(n-2)\,\mathrm{N}'_0 + (-)^{n-2}(n-1) = 1.$

Hence

$$\begin{aligned} 1 &= \mathrm{N}_{n-3} - \mathrm{N}_{n-4} + \ldots + (-)^{n-3}\mathrm{N}_0 + (-1)^{n-2} \\ &\quad - 2\mathrm{N}_{n-4} + \ldots + (-)^{n-3}\,2\mathrm{N}_0 + 2\,(-1)^{n-2} \ldots \\ &= \mathrm{N}_{n-3} - 3\mathrm{N}_{n-4} + 6\mathrm{N}_{n-5} - \ldots + (-)^{n-3}{}_{n-1}\mathrm{C}_{n-3}\,\mathrm{N}_0 + (-)^{n-2}{}_n\mathrm{C}_{n-2}, \end{aligned}$$

i.e.,

$$ {}_n\phi_{n-2}(\mathrm{N}) = 1. \tag{4.53}$$

We may call the pyramid in this case a pyramid of the second order. A pyramid of order r is one whose base is a pyramid of order $r-1$. In two dimensions a pyramid of first order is a triangle, in three dimensions a pyramid of second order is a tetrahedron, and generally in n dimensions a pyramid of order $(n-1)$ is a simplex.

It can be shown generally by induction that for a pyramid of order r we have the $r+1$ relations

$$ {}_n\phi_s(\mathrm{N}) = 1, \quad (s = n, n-1, \ldots, n-r). \tag{4.54}$$

§ 5. *Relations between the volume and the angle-sums of a polytope bounded entirely by simplexes.*

5.1. Take any point in the interior of the polytope, and join it to all the vertices, thus dividing it centrally into simplexes. Let Σ_r denote the sum of the angular regions at the r-dimensional edges for the polytope, and Σ'_r the corresponding sum for a constituent simplex; V the whole volume and V′ that of a constituent simplex.

Then

$$\begin{aligned}
\Sigma\Sigma'_n &= N_{n-1}\Sigma_n,\\
\Sigma\Sigma'_{n-1} &= \Sigma_{n-1} + N_{n-2}\Sigma_n,\\
\Sigma\Sigma'_{n-2} &= \Sigma_{n-2} + N_{n-3}\Sigma_n,\\
&\cdots\cdots\cdots\cdots\\
\Sigma\Sigma'_1 &= \Sigma_1 + N_0\Sigma_n,\\
\Sigma\Sigma'_0 &= \Sigma_0 + \Sigma_n,\\
\Sigma V' &= V.
\end{aligned}$$

5.2. For each simplex we have the relations

$$_{n+1}\phi_r(\Sigma') = {}_{n+1}\phi_{n-r+1}(\Sigma').$$

Summing these equations and substituting for $\Sigma\Sigma'_r$ we have for $r = 1, 2, \ldots [\frac{1}{2}(n-1)]$

$$\begin{aligned}
&{}_{n+1}\phi_r(\Sigma) + (N_{r-2} - {}_{n-r+2}C_1 N_{r-3} + \ldots + (-)^{r-1}\,{}_nC_{r-1})\Sigma_n\\
&\quad = {}_{n+1}\phi_{n-r+1}(\Sigma) + (N_{n-r-1} - {}_{r+1}C_1 N_{n-r-2} + \ldots + (-)^{n-r}\,{}_nC_{n-r})\Sigma_n,
\end{aligned}$$

that is,

$$\begin{aligned}
{}_{n+1}\phi_r(\Sigma) - {}_{n+1}\phi_{n-r+1}(\Sigma) &= \{{}_n\phi_{n-r}(N) - {}_n\phi_{r-1}(N)\}\Sigma_n\\
&= \{{}_n\phi_r(N) - {}_n\phi_{r-1}(N)\}\Sigma_n \qquad \text{by (4.42)}\\
&= {}_{n+1}\phi_r(N)\,.\,\Sigma_n = -\,{}_{n+1}\phi_{n-r+1}(N)\,.\,\Sigma_n, \qquad (5.21)
\end{aligned}$$

since ${}_nC_r + {}_nC_{r-1} = {}_{n+1}C_r$.

But for $r = 0$

$$\begin{aligned}
{}_{n+1}\phi_0(\Sigma) - {}_{n+1}\phi_{n+1}(\Sigma) &= \{(N_{n-1} - 1) - N_{n-2} + \ldots + (-)^{n-1}N_0 + (-1)^n\}\Sigma_n\\
&= 0, \qquad (5.22)
\end{aligned}$$

i.e., the first relation, viz.,

$$\Sigma_n - \Sigma_{n-1} + \Sigma_{n-2} - \ldots + (-)^n\Sigma_0 + (-)^{n+1}V = V,$$

is the same for all polytopes bounded entirely by simplexes.

We have also the relation,

$$2\Sigma_{n-1} = N_{n-1}\,.\,\Sigma_n. \qquad (5.23)$$

The relations between the volume and the angle-sums in radian measure for $n = 2, 3, 4$ are

$$n = 2: \quad V = k^2\{S_0 - (N_1 - 2)\pi\}, \qquad (5.24)$$

$$n = 3: \quad 0 = S_0 - S_1 + (N_2 - 2)2\pi, \qquad (5.25)$$

$$n = 4: \quad V = \tfrac{2}{3}k^4\{S_2 - S_1 + S_0 + (2 - N_3)\pi^2\}, \qquad (5.26)$$

$$V = \tfrac{2}{15}k^4\{2S_2 - 3S_1 + 5S_0 - (2N_0 + N_3 - 10)\pi^2\}.$$

§ 6. *Relations between the volume and the angle-sums for any Eulerian polytope.*

6.1. Take any r-dimensional boundary and divide it centrally. Then if α_r

is the volume of the angular region at that boundary as edge, we have, if Σ_s is the sum of the angular regions at an s-dimensional edge for the given polytope, and Σ'_s that for the transformed polytope,

$$\Sigma'_s = \Sigma_s \qquad (s = n, n-1, \ldots, r+1)$$
$$\Sigma'_r = \Sigma_r + (N_{r-1,r} - 1)\alpha_r,$$
$$\Sigma'_s = \Sigma_s + N_{s-1,r}\alpha_r \qquad (s = r-1, r-2, \ldots, 1)$$
$$\Sigma'_0 = \Sigma_0 + \alpha_r,$$
$$V' = V.$$

Then

$$\begin{aligned}\Sigma'_n - \Sigma'_{n-1} + \Sigma'_{n-2} - \ldots + (-)^{n-1}\Sigma'_1 + (-)^n\Sigma_0' + (-)^{n+1}V' \\ = \Sigma_n - \Sigma_{n-1} + \ldots + (-)^{n-1}\Sigma_1 + (-)^n\Sigma_0 + (-)^{n+1}V \\ + (-)^{n-r}\{(N_{r-1,r} - 1) - N_{r-2,r} + \ldots + (-)^{r-1}N_{0,r} + (-1)^r\}\alpha_r \\ = \Sigma_n - \Sigma_{n-1} + \ldots + (-)^n\Sigma_0 + (-)^{n+1}V,\end{aligned}$$

i.e.,

$$_{n+1}\phi_{n+1}(\Sigma') = {}_{n+1}\phi_{n+1}(\Sigma).$$

Thus the function $_{n+1}\phi_{n+1}(\Sigma)$ is not altered if all the boundaries are divided centrally. If all the boundaries of all dimensions are divided centrally the polytope is transformed into one bounded entirely by simplexes, hence *the relation*

$$\Sigma_n - \Sigma_{n-1} + \Sigma_{n-2} - \ldots + (-)^n\Sigma_0 + (-)^{n+1}V = V \qquad (6.11)$$

is true for all Eulerian polytopes.

6.2. For the other relations it is found that the angles at the individual boundaries are involved. We shall work out as an example the case of $n = 4$.

Consider a polytope in space of 4 dimensions. Let N_r be the number of its r-dimensional boundaries, N_{pq} the number of p-dimensional boundaries of, or incident with, a particular q-dimensional boundary.

First choose any 2-dimensional boundary and divide it centrally. Let Σ_r be the sum of the angular regions for the polytope, Σ'_r those for the transformed polytope.

Then

$$\Sigma'_4 = \Sigma_4,$$
$$\Sigma'_3 = \Sigma_3,$$
$$\Sigma'_2 = \Sigma_2 + (N_{12} - 1)\alpha_2,$$
$$\Sigma'_1 = \Sigma_1 + N_{02}\alpha_2,$$
$$\Sigma'_0 = \Sigma_0 + \alpha_2,$$
$$V' = V,$$

where α_2 is the angle at the 2-dimensional boundary. Also $N_{02} = N_{12}$.

Let the function

$$ {}_5\phi_4(\Sigma) - {}_5\phi_1(\Sigma) \equiv \Sigma_3 - 2\Sigma_2 + 3\Sigma_1 - 5\Sigma_0 + 10V \equiv \psi(\Sigma). $$

Then

$$ \begin{aligned} \psi(\Sigma') &= \Sigma'_3 - 2\Sigma'_2 + 3\Sigma'_1 - 5\Sigma'_0 + 10V' \\ &= \psi(\Sigma) + (-2N_{12} + 2 + 3N_{02} - 5)\,\alpha_2 \\ &= \psi(\Sigma) + (N_{12} - 3)\,\alpha_2. \end{aligned} $$

Hence if all the 2-dimensional boundaries are divided centrally

$$ \psi(\Sigma') = \psi(\Sigma) + \Sigma(N_{12} - 3)\,\alpha_2. \qquad (6.21) $$

Next choose any 3-dimensional boundary and divide it centrally. Let N'_{p3} be the number of its boundaries of p dimensions, Σ''_r the sum of the angular regions at an r-dimensional edge for the polytope thus further transformed.

Then

$$ \begin{aligned} \Sigma''_4 &= \Sigma'_4, \\ \Sigma''_3 &= \Sigma'_3 + (N'_{23} - 1)\,\alpha_3, \\ \Sigma''_2 &= \Sigma'_2 + N'_{13}\alpha_3, \\ \Sigma''_1 &= \Sigma'_1 + N'_{03}\alpha_3, \\ \Sigma''_0 &= \Sigma'_0 + \alpha_3, \\ V'' &= V', \end{aligned} $$

where α_3 is the angle at the 3-dimensional boundary, $= \frac{1}{2}\Sigma_4$.

Then $\psi(\Sigma'') = \psi(\Sigma') + (N'_{23} - 1 - 2N'_{13} + 3N'_{03} - 5)\,\frac{1}{2}\Sigma_4$, and when all the 3-dimensional boundaries are divided centrally

$$ \psi(\Sigma'') = \psi(\Sigma') + \Sigma(N'_{23} - 2N'_{13} + 3N'_{03} - 6)\,\tfrac{1}{2}\Sigma_4. \qquad (6.22) $$

But $N'_{03} = N_{03} + N_{23}$,

$N'_{13} = N_{13} + \Sigma N_{12}$ (the summation extending over the faces of the boundary)

$= N_{13} + 2N_{13} = 3N_{13}$,

$N'_{23} = \Sigma N_{12} = 2N_{13}$.

Hence

$$ \begin{aligned} N'_{23} - 2N'_{13} + 3N'_{03} - 6 &= 3N_{23} - 4N_{13} + 3N_{03} - 6 \\ &= -N_{13}, \end{aligned} $$

and, since the polytope is now bounded entirely by simplexes, by (5.21)

$$ {}_5\phi_4(\Sigma'') - {}_5\phi_1(\Sigma'') = -{}_5\phi_1(N'')\,.\,\Sigma_4, $$

i.e.,

$$ \psi(\Sigma'') = (5 - N''_0)\,\Sigma_4. $$

But

$$N''_0 = N'_0 + N'_3,$$
$$N'_0 = N_0 + N_2, \qquad N'_3 = N_3.$$

Therefore

$$N''_0 = N_0 + N_2 + N_3.$$

Hence finally

$$\begin{aligned}\psi(\Sigma) &= \psi(\Sigma') - \Sigma(N_{12} - 3)\alpha_2 \\ &= (5 - N''_0)\Sigma_4 + \tfrac{1}{2}\Sigma_4 \cdot \Sigma N_{13} - \Sigma(N_{12} - 3)\alpha_2 \\ &= (5 - N_0 - N_2 - N_3 + \tfrac{1}{2}\Sigma N_{13})\Sigma_4 - \Sigma(N_{12} - 3)\alpha_2 \qquad (6.23)\end{aligned}$$

6.3. For the regular polytopes, or more generally the homogeneous polytopes, these relations become simplified since N_{12} and N_{13} are the same for all the boundaries.

We have
$$N_{13} = N_{23} + N_{03} - 2,$$
$$\Sigma N_{23} = \Sigma N_{32} = N_2 N_{32} = 2N_2,$$
$$\Sigma N_{03} = \Sigma N_{30} = N_0 N_{30},$$
hence
$$\Sigma N_{13} = 2N_2 + N_{30} N_0 - 2N_3.$$
Also $\Sigma(N_{12} - 3)\alpha_2 = (N_{12} - 3)\Sigma_2$, and $\Sigma_3 = \frac{1}{2}N_3 \cdot \Sigma_4$.

Hence we have

$$\{\tfrac{5}{2}N_3 + (1 - \tfrac{1}{2}N_{30})N_0 - 5\}\Sigma_4 + (N_{12} - 5)\Sigma_2 + 3\Sigma_1 - 5\Sigma_0 + 10V = 0, \quad (6.31)$$

or in terms of the radian measures

$$2k^4\{[5N_3 + (2 - N_{30})N_0 - 10]\pi^2 + (N_{12} - 5)S_2 + 3S_1 - 5S_0\} + 15V = 0. \quad (6.32)$$

We have also by (6.11)

$$2k^4\{(2 - N_3)\pi^2 + S_2 - S_1 + S_0\} = 3V. \quad (6.33)$$

Eliminating V we get

$$N_{12}S_2 - 2S_1 = (N_{30} - 2)N_0\pi^2. \quad (6.34)$$

6.4. In particular for the regular polytopes, if α_2, α_1, α_0 are the radian measures of the angles at a face, an edge, and a vertex, $S_2 = N_2\alpha_2$, $S_1 = N_1\alpha_1$, $S_0 = N_0\alpha_0$. Equation (6.34) then becomes

$$N_{02}N_2\alpha_2 - 2N_1\alpha_1 = (N_{30} - 2)N_0\pi^2;$$

but $N_{02}N_2 = N_{20}N_0$ and $2N_1 = N_{01}N_1 = N_{10}N_0$,
hence

$$N_{20}\alpha_2 - N_{10}\alpha_1 = (N_{30} - 2)\pi^2. \quad (6.41)$$

This result is easily verified directly by drawing a hypersphere round a vertex and applying equation (5.25), which gives

$$S'_1 - S'_2 + (N_{30} - 2)\pi^2 = 0,$$

S'_1 and S'_2 being the sums of the angles at the edges and faces at that vertex, and π^2 (for 4 dimensions) replacing 2π (for 3 dimensions), half the complete angle at a point. Then $S'_1 = N_{10}\alpha_1$ and $S'_2 = N_{20}\alpha_2$.

Since $N_{30} - N_{20} + N_{10} = 2$, equation (6.41) may also be written

$$N_{20}(\alpha_2 - \pi^2) = N_{10}(\alpha_1 - \pi^2). \tag{6.42}$$

If ${}_rN_{pq}$ denotes the number of p-dimensional boundaries through a q-dimensional boundary and lying in an r-dimensional boundary

$$N_{32}N_{20} = {}_3N_{20}N_{30} = {}_3N_{10}N_{30} = N_{31}N_{10},$$

and $N_{32} = 2$, hence $2N_{20} = N_{31}N_{10}$, and (6.42) may be further simplified to

$$N_{31}(\alpha_2 - \pi^2) = 2(\alpha_1 - \pi^2). \tag{6.43}$$

For the 5-, 8-, 24- and 120-cells $N_{31} = 3$, for the 16-cell $N_{31} = 4$, and for the 600-cell $N_{31} = 5$.

The numbers $N_{01} = N_{12} = k_1$, ${}_3N_{10} = {}_3N_{20} = k_2$, and $N_{21} = N_{31} = k_3$ are the *fundamental numbers* of the regular polytopes,* in terms of which all the numbers N_{pq} and the ratios of the numbers N_r can be expressed.

* See the author's paper "The regular divisions of space of n dimensions and their metrical constants," Palermo, 'Rend. Circ. mat.,' vol. 48, pp. 1–14 (1924).

Thermal Changes in Iron-Manganese Alloys, Low in Carbon.

By Sir Robert Hadfield, F.R.S.

(Received March 17, 1927.)

Introduction.

As this paper deals with one part only of a very extended research on iron-manganese alloys low in carbon produced by the author, an outline of the general results obtained would appear useful and appropriate as an introduction.

Although the peculiar qualities of the alloy known as manganese steel, in which the carbon present is high, have been examined from many points of view, in the present research the author's special object has been to study this series with the carbon content reduced to so low a percentage as to be practically negligible.

The alloys have been submitted to a very wide range of tests, numbering over eight hundred—mechanical, physical, magnetic, chemical, and metallographic—the research forming a continuation of the author's original work, which resulted in the discovery of manganese steel, the now well-known alloy which contains 12 to 14 per cent. of manganese, with high carbon; that is, 1·00 to 1·25 per cent. This material, after "heat treatment and quenching," possesses very high tenacity combined with extraordinary toughness.

As a result of this research the author has found that the low carbon alloys here described when heat-treated and quenched do *not* possess the remarkable combination of tenacity and ductility displayed by manganese steel in anything like the same degree. At the time of the original research, some forty years ago, it happened, fortunately for industrial applications, that the only known and accessible materials with which to produce his iron-manganese alloys were spiegel and ferro-manganese, each containing comparatively high percentages of carbon. Therefore, the alloys so produced were high in carbon. Metallic manganese of the purity that would have produced a low-carbon alloy was not then available.

The important fact has been proved by this research that the whole range of these low-carbon alloys is much less amenable to the effects of this heating and quenching, a fact of primary metallurgical importance.

In respect of mechanical qualities the alloys* group themselves into five ranges of composition :—

* For detailed composition *v. infra* under fig. 1.

(1) Mn under 4 per cent.—alloys comparatively soft and tough, becoming harder and less tough with increasing Mn.

(2) Mn 4–10 „ „ characterised by comparatively high Brinell hardness (region of 400) and brittleness.

(3) Mn 10–15 „ „ a transition zone in which the ductility improves and hardness decreases as Mn increases.

(4) Mn 15–39 „ „ character approaching manganese steel—Brinell hardness about 200, tenacity and ductility considerable, with a fair capacity for work-hardening.

(5) Mn 61·50 „ „ showed intermediate qualities—the product could not be forged successfully at any temperature between 800 and 1100° C.

(6) Mn about 80 „ „ Hard and brittle.

Under heat treatment some changes occur in the mechanical qualities of these alloys, but in no case is anything like a transformation effected, the essential characteristics remain.

Manganese steel heat-treated and water-quenched, possesses great tenacity, namely, about 70 tons per square inch, with about 70 per cent. elongation and comparatively low Brinell hardness. Annealing for a considerable time works a complete transformation, the steel becoming extremely hard and brittle—that is, up to Brinell hardness 500. At the same time, the micro-structure, from being of the purely austenitic type, now shows carbide deposited in considerable amounts from solid solution in acicular formation. Important magnetic changes also follow through the whole range of the alloys.

Thus it is seen that heat-treatment operates mainly through the carbon present, and its physical and chemical relations with the iron and manganese. It is further interesting to find that the characteristic ductility and toughness of manganese steel are almost completely lost in these alloys in which the carbon is practically eliminated. Thus, as is shown elsewhere,* specimen 1379/E shows that the elongation is reduced to no more than 2 per cent. in the practically carbon-free material, as against about 70 per cent. for ordinary manganese steel with high carbon, an indication of the important rôle played by the carbon.

In connection with the brittleness shown in Group 2 above, the following

* *Vide* Hadfield, "Alloys of Iron and Manganese containing Low Carbon," Table VI, 'Journ. Iron and Steel Inst.,' 1927 (*in the press*).

points are also of interest. In the author's original manganese steel (see 'Proc. Inst. Civ. Eng.,' 1888), with a C : Mn ratio of 1 : 10, a somewhat similar combination of brittleness and hardness was noted. The alloy in question (0·48 C, 5·00 Mn) was so brittle that it could be readily pounded with a hand-hammer, and thus seems to be connected with the similar brittleness above-mentioned, although found there in a less-marked degree. This quality in the alloys of comparatively low Mn percentage is evidently derived mainly from the Fe-Mn relation, although the presence of carbon accentuates the brittleness and increases the hardness to over 600 Brinell.

In determining magnetic properties three conditions were examined—forged, annealed, water-quenched. In each case also a bar was water-quenched from 1000° C., and afterwards reheated to 500° C. for 60 hours, which treatment confers strong magnetic properties upon manganese steeel. (The reasons for applying this special annealing are given in the author's joint paper with the late Prof. Onnes and Dr. Woltjer, published in the 'Roy. Soc. Proc.,' A, vol. 99, 1921.) Alloys of the groups with high Mn percentage (more than 16 per cent.) became almost completely non-magnetic, whether of low carbon content or with 1 to 1·30 per cent., as in ordinary manganese steel.

The author gives a table* of the magnetic results, together with the electrical resistance of the alloys in water-quenched condition. This shows that the magnetism falls away in two stages—a fairly slow drop from 0 to 7 per cent. Mn, and a more rapid drop from 7 per cent. to about 16 per cent. It is further shown in the present research that the non-magnetic quality persists not only up to a manganese content of 39 per cent., but also that the 60 and 80 per cent. alloys possess this characteristic, as is the case with the metal manganese itself unalloyed with iron.

With the specimens reheated to 500° C. for 60 hours, it is seen that their specific magnetism remains in all cases practically the same as in the original water-quenched condition. The fact noted in the joint paper above mentioned, that these Fe-Mn alloys, in contradistinction to manganese steel, are not affected in their magnetism by such heat-treatment is fully confirmed.

In view of the fact that manganese steel, containing about 13 per cent. Mn and 1·25 per cent. C, is practically non-magnetic, it is somewhat surprising to find that removal of the carbon produced such a great change; in fact, a complete transformation in this characteristic—that is, an alloy containing 12·95 per cent. manganese and 0·09 per cent. carbon, shows no less than 12 per cent. specific magnetism. On the other hand, a low-carbon iron-

* *Vide* Hadfield, *loc. cit.*, Table XVI.

manganese alloy with 17 per cent. manganese and 0·15 per cent. carbon, shows absence of magnetic susceptibility.

The author has prepared a table* of comparative magnetic data for low-carbon alloys, and for those containing about 1 per cent. carbon. Those containing the higher percentages of manganese are non-magnetic, even in the practical absence of carbon, but the presence of carbon causes the non-magnetic quality to appear at a manganese percentage much lower than in its absence. With about 1 per cent. carbon not more than 7 per cent. manganese is necessary; while with low-carbon 16 to 17 per cent. or more of manganese is requisite to suppress the magnetic qualities of the iron.

It is thus seen to be an important conclusion from this research that the non-magnetic qualities and want of susceptibilty of these Fe-Mn alloys are largely influenced by the proportion of carbon present.

Many other interesting results of this extensive research must necessarily remain unrecited in this brief summary. It is hoped, however, that this introduction will serve to give some idea of the general nature and peculiarities of the alloys, the thermal changes of which form the subject of the following paper.

1.—The present research into the physical changes which these alloys undergo during heating and cooling, forms an extension of this work.

From the fact that alloys of iron and nickel, which at about 30 per cent. nickel become practically non-magnetic, can be transformed and rendered magnetic by cooling to a sufficiently low temperature, it had been inferred that iron-manganese alloys were similar in character and should behave in the same manner. This theory was also in fact extended to manganese steel, the alloy containing between 12 and 14 per cent. manganese and, in addition, carbon to the extent of about 1 to 1¼ per cent., that is, it was supposed that this non-magnetic alloy, if cooled to a sufficiently low temperature, would become magnetic.

The author's various researches, and especially that carried out in collaboration with the late Professor Kamerlingh Onnes and Dr. H. R. Woltjer, recorded in the joint paper read before this Society in 1921, has definitely shown that no such transformation occurs in the non-magnetic alloys of iron and manganese even down to temperatures as low as that of liquid helium, namely,

* *Vide* Hadfield, *loc. cit.*, Table XIII.

—269° C. It is thus established with reasonable certainty that there is an essential difference in metallurgical character between the alloys of iron and manganese and those of iron and nickel. Whereas the non-magnetic alloys of iron and nickel undergo critical magnetic changes at a temperature below atmospheric, it is found that no such changes occur in the iron-manganese alloys.

It was desirable, however, to trace the actual progress of the thermal changes in their position on the temperature scale with varying manganese percentage so as to ascertain whether the temperature of the magnetic transformation does in fact pass below atmospheric temperature. The complete series of alloys from 1·70 to 38·90 per cent. manganese, also an additional alloy containing 83·50 per cent. manganese, described in the author's recent paper to this Society, presented an excellent opportunity for obtaining this information. The determination of the temperature of the critical changes in the iron-manganese alloys has therefore been the object of the present research.

The temperatures at which the magnetic transformations take place in such alloys up to about 13 per cent. manganese have to some extent previously been studied. On the other hand, as the non-magnetic alloys of iron and manganese with low carbon are not reached until about 16 per cent. of manganese is exceeded, from the present point of view the required information has not before been provided.

It will be seen from the results of the present research that magnetic transformations are confined to the range of magnetic alloys with less than 16 per cent. manganese. The non-magnetic alloys above 16 per cent. appear to be entirely devoid of magnetic transformations, no matter how heat-treated, or even when cooled to the lowest known temperatures. These iron-manganese alloys are therefore magnetically inert, being in this respect similar to metals such as copper. It should be remembered, too, that these alloys, being practically free from carbon, do not come under the term "manganese steel."

In the author's previous joint paper, already referred to, and in which these iron-manganese alloys are dealt with at liquid hydrogen and helium temperatures, it was suggested that the non-magnetic qualities found in these alloys with higher percentages might be due to actual combination of the iron and manganese in much the same way as, for example, the magnetism of iron is suppressed when chemically combined with oxygen as ferrous oxide. From this point of view the iron-manganese compound would probably be represented by the formula Fe_5Mn, containing 16·40 per cent. manganese, which is very close to the percentage at which non-magnetic properties appear. Thus the

non-magnetic alloys of iron and manganese, in which carbon is practically absent, would in such cases consist of this non-magnetic compound, with a more or less additional amount of the non-magnetic metal manganese, according to the total amount of manganese content. The absence of magnetic transformations in these non-magnetic alloys would seem to give further support to such a theory.

The present research has also disclosed certain peculiar features in these iron-manganese alloys in a curious doubling of the magnetic transformation on cooling, though not on heating; that is, for certain of the magnetic alloys the first recovery of magnetism which takes place on cooling only exists over a limited range of temperature, the alloy becoming again practically non-magnetic, until finally, with still further reduction in the temperature, the full magnetic properties reappear and remain down to atmospheric temperature.

2.—*Method of Investigation.*

Some discussion regarding the critical ranges of iron-manganese alloys was presented in the joint paper above referred to by Onnes, Woltjer and the author. This, however, had more specific reference to the magnetic transformations, and was directly concerned with the explanation as to why the alloys containing about 16 per cent. manganese and above are non-magnetic, and so far as the author is aware, their critical temperatures over the whole range covered by the present series of alloys have not previously been determined.

The author has therefore prepared the heating and cooling curves shown in figs. 1 and 2 of the inverse rate type. These were made on selected specimens representing stages in manganese percentage up to 17·10 per cent. manganese and in their forged condition; a similar determination was also made on 13791 containing 83·50 per cent. manganese, the series being eventually completed by the addition of material J with 61·50 per cent. manganese. Both these high percentage alloys were necessarily examined in their cast condition. In addition to these heating and cooling curves showing thermal changes, observations, shown on the same diagrams, have been taken of the temperatures at which loss and regain of magnetism occur—that is, for those specimens which are magnetic at ordinary temperatures.

To obtain the inverse rate curves described in this paper the author employed the usual methods, that is, similar to those described in his paper on "Heating and Cooling Curves of Manganese Steel" (Iron and Steel Institute, 1913), with the exception that an electric furnace was used instead of a gas furnace.

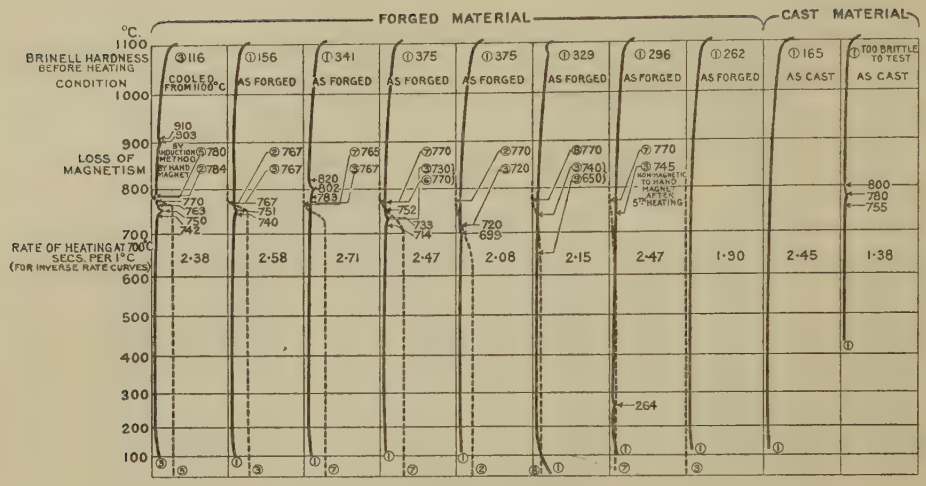


FIG. 1.—Heating Curves.

ALLOY No.	1727D	1379A	1379B	1379C2	1379D	1379E	1379E3	1379G	1379J	1379I
C	0·11	0·07	0·06	0·08	0·07	0·09	0·14	0·15	0·24	0·29
Si	0·20	0·03	0·09	0·08	0·10	0·12	0·26	0·35	1·14	1·36
S	0·049	0·120	0·098	0·076	0·084	0·103	0·051	0·060	0·062	0·054
P	0·050	0·070	0·068	0·064	0·059	0·050	0·044	0·051	0·039	0·050
Mn	0·06	1·68	3·95	6·75	9·45	12·95	14·30	17·10	61·50	83·50

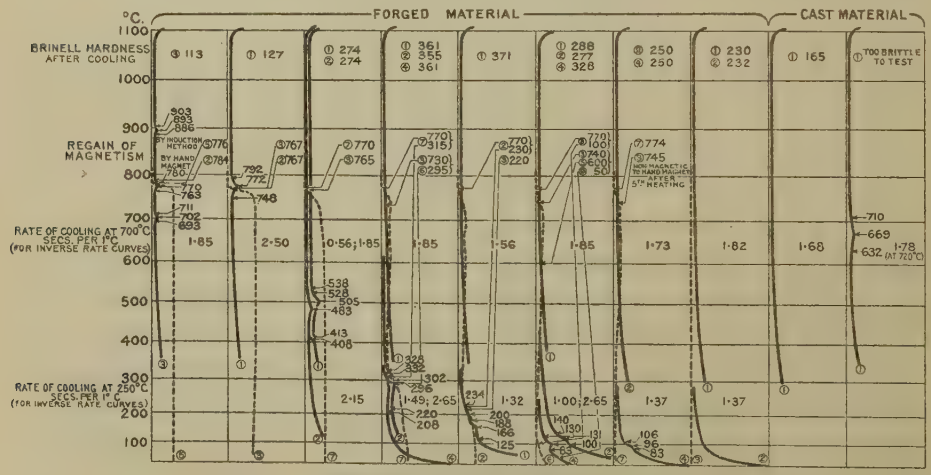


FIG. 2.—Cooling Curves.

Showing absorptions and evolutions of heat, also magnetic changes, during heating and cooling of iron-manganese alloys, low in carbon.

Curves shown by FULL LINES were taken by the Inverse Rate Method and indicate THERMAL CHANGES. Curves shown by *broken lines* were taken by an Electrical Induction Method and indicate the *Magnetic Changes*: the abscissa of the latter curves at any temperature is proportional to the Magnetism of the specimen in a field of 15 c.g.s.

Numbers ringed indicate the serial number of the heating subsequent to forging, or in the case of specimens I and J subsequent to casting.

A uniform maximum temperature of 1100° C. (actually, 1098° to 1107°) was adopted, and the specimens were each heated under the same conditions so as to be comparative. In cooling, the presence of critical points at relatively low temperatures in the alloys with the higher percentages of manganese rendered it necessary to take a second cooling curve at a most rapid rate, so that the cooling was not unduly slow at the lower temperatures, with risk of masking to some extent the critical points.

For the purpose of determining the magnetic transformations a separate heating of the specimens was made, inside a refractory tube wound with a magnetising coil of non-magnetic nickel-chromium alloy wire, the whole being placed in an electric furnace. A magnetising field of 15 c.g.s. was employed.

To observe the magnetic condition of the specimen, it was surrounded by a secondary coil of the same wire, wound on a silica tube, this tube being inside that supporting the primary winding. Side by side with this secondary coil was a similar one, and both were connected in series, but magnetically opposed, with a ballistic galvanometer. With no magnetic material in the furnace, making or breaking the magnetising current produced no kick in the galvanometer, whether the furnace heating current was on or off. With the specimen inside one of the secondary coils, the magnitude of the kick on the galvanometer, when the magnetising circuit was opened or closed, was therefore proportional to its magnetic induction for a field of 15 c.g.s. In this way, during the heating and cooling of the specimens, a series of observations was made and plotted, resulting in the curves shown in the diagrams. For each material the temperature of loss and regain of magnetism was in this way determined with sufficient clearness.

A further indication was obtained by applying a permanent magnet to the specimens during heating and cooling. With practice the loss of magnetism on heating, and its return on cooling, could readily be detected by the pull of the magnet. The indications obtained in this way are obviously not so reliable as those determined by the induction method. In general, the temperatures indicated by the magnet are rather lower, no doubt owing to a larger degree of magnetisation being necessary to affect it. For the determination of the induction curves the rate of heating and cooling was five seconds per 1° C., and for the hand magnet tests, ten seconds per 1° C. The maximum temperature in both cases was about 810° C.

3.—*Experimental Data and Discussion of Results.*

The diagrams, figs. 1 and 2, show for all the magnetic alloys, that is, up to 14·30 per cent. manganese, a reversible magnetic transformation, at practically the same temperature as the A_2 magnetic change in iron. The temperature at which this change occurs in the present alloys on heating is from 765° to 780°, and on cooling from 767° to 776°.

As regards heating, this one critical change is all that is noticeable, and this applies also to the cooling for the alloys up to 3·95 per cent. manganese. For these the magnetism, after a slight fall of temperature below the critical point, reaches practically its full value and remains more or less constant down to room temperature. With the other magnetic alloys of 6·75 per cent. manganese and upwards to 14·30 per cent., however, on cooling below the reversible magnetic change point the magnetism after reaching a maximum falls off again with decreasing temperature until a second critical temperature is reached, at which there is a further recovery. This recovery is maintained and increased until, when room temperature is reached, the specimen has regained approximately its original magnetic value. The magnetic changes in these alloys are naturally smaller in magnitude generally than for the more magnetic alloys with less manganese, and diminish in strength as the manganese content increases.

This second and lower critical magnetic change point appears in alloy C/2 (6·75 per cent. manganese) at 315°, and is progressively lowered with increase in manganese to about 100° in alloy E (12·95 per cent. manganese), becoming at the same time much weaker. In alloy E/3 (14·30 per cent. manganese) no such critical point is discernible. This is only to be expected in view of the feebly magnetic qualities of the material. It is interesting to note that this material E/3, although never displaying magnetism to a high degree, is more strongly magnetic at about 600° than at room temperature.

Specimen G (17·10 per cent. manganese), as on heating, gives no indication of magnetism at any temperature on cooling.

A noticeable feature in these magnetic determinations was the fact that this lower magnetic change point was not fixed, but in all cases on repetition appeared at a rather lower temperature. Thus in E (12·95 per cent. manganese) during the cooling following the seventh heating it appeared at 150°. In the eighth cooling the temperature was 100°. Check tests on specimen D (9·45 per cent. manganese) with the hand-magnet gave 220° and 205° in the third and fourth coolings respectively, while with C/2 (6·75 per cent. manganese) the temperatures were 300° and 295° in the fifth and sixth coolings.

For this reason the serial number of each heating and cooling is indicated against all the curves in the diagrams. Whether such alteration is due to a physical change caused by repeated heating, or to a reduction by oxidation in the already small carbon content, has not yet been investigated. No such variation occurs in the position of the upper magnetic transformation, which remains fixed at closely the same temperature with repeated heating. Thus with specimen E (12·95 per cent. manganese) in the eighth cooling, in which the lower magnetic change had shifted through 50°, the upper magnetic change was found at 770°, as in the seventh cooling.

It will now be interesting to examine how far these magnetic changes are associated with thermal evolutions shown by the inverse rate curves (fig. 1). For alloys A and B only one point appears clearly, but a slight though definite protuberance seen on the curves seems to indicate a further thermal transformation in each material. The upper of these two points seems to be the Ac_3 change progressively lowered in temperature from 903° to 802° with increase of manganese from 0·06 to 3·95 per cent. The lower point is evidently associated with the Ac_2 change agreeing reasonably well in temperature with the observed loss of magnetism.

With further increase in manganese to 6·75 and 9·45 per cent. only one absorption of heat is observed, ending at 752° and 730° respectively, as might be expected from the further lowering of Ac_3, thus bringing down with it Ac_2.

With 12·95 and 14·30 per cent. manganese, although these alloys have magnetic transformations, no corresponding indications are seen in the inverse rate curves. The latter, however, displays a critical point at as low as 264°. On repetition a similar point was found at 230°.

In the inverse rate curves for cooling (fig. 2), it is curious to find that none of these except that for A (1·68 per cent. manganese) gives any indication corresponding to the upper (reversible) magnetic change. The lower magnetic change seen in the magnetic alloys from 6·75 per cent. manganese upwards is, on the other hand, associated with an evolution of heat, diminishing in intensity with increase of manganese. This evolution is discernible up to 14·30 per cent. manganese, although the magnetic change could not be detected above 12·95 per cent.

The curve for D (9·45 per cent. manganese) shows a curious doubling of the heat evolution which is also slightly but definitely reflected in the curve of magnetism. Generally this examination of the magnetic alloys of this series leads to the following conclusion. The Ac_3 point, while being lowered as the result of increasing manganese, in due course reaches the temperature

(about 770° C.) of the Ac_2 magnetic change, and possibly brings down with it the absorption or evolution of heat associated with the latter change. It, however, leaves undisturbed the actual magnetic change, which remains reversible. On cooling, and where Ar_3 has passed below 770° C. it is accompanied by a separate magnetic transformation; the magnetism which has been restored at 770° C. and has fallen away again with lowering temperature, returns more or less completely at Ar_3. Curiously no corresponding change takes place on heating, that is Ac_3 is devoid of magnetic change while it is below 770° C.

In this connection, however, the behaviour of alloy B (3·95 per cent. manganese) should be specially noted. Its single evolution of heat at about 500° C. on cooling is unaccompanied by any noticeable magnetic change. This alloy is distinguished from the alloys of higher percentage in that although Ar_3 is below 770° C., Ac_3 is above the magnetic change on heating; although this does not supply an obvious explanation for its different behaviour, it may be that for a separate magnetic change to be associated with Ar_3 on cooling, the Ac_3 change should precede the magnetic change on heating. The magnetic behaviour of these alloys of iron and manganese during heating and cooling is, in fact, so interesting and unusual as to justify further investigation.

On the general question of the non-magnetic qualities found in the alloys of iron and manganese of the higher percentages, it seems clear from the present tests and from the magnetic examination recorded in the previous section, that the loss of magnetism which occurs when the manganese reaches about 16 per cent. is the culminating point of a progressive reduction in the magnetic quality with increase of manganese—not merely as found at ordinary temperature, but at all temperatures.

It has been too readily assumed by some writers that the magnetic transformation, in its downward course with increasing manganese, passes below atmospheric temperature; hence the alloys of higher percentage above where this occurs are naturally found at ordinary temperatures to be non-magnetic. On the contrary, as was shown in the joint research referred to by Onnes, Woltjer and the present author, this explanation cannot be correct, since no evidence of transformations below ordinary temperature was found, that is, for iron manganese alloys. With iron nickel alloys, on the other hand, magnetic transformations did occur with immersion in liquid air.

In the present research it seems clear that the magnetic strength of the alloys becomes progressively weaker with increase of manganese, and the final transformation also as a consequence diminishes in intensity, finally vanishing

before 17·10 per cent. manganese is reached, and also before this transformation has reached ordinary temperature.

It remains to refer to the alloys of specially high percentage J and I. Like G (17·10 per cent. manganese) which is also non-magnetic, J (61·50 per cent.) is devoid of critical points in its inverse rate curves, and it was not thought worth while to examine it for magnetic changes which are absent from G. In I (83·50 per cent. manganese) appear corresponding critical points on its heating and cooling curves at 800° and 710° respectively, apparently quite unconnected with those observed in the alloys containing the lower percentages of manganese. The actual nature of the change represented by these critical points has not so far been examined, though possibly it has its origin in some transformation occurring in the metal manganese itself.

4.—*Conclusion.*

The author trusts that the data obtained in this research regarding the thermal changes in the alloys of iron and manganese will be found helpful towards explaining their special characteristics.

The suggestion that the non-magnetic properties of the alloys of iron and manganese of the higher percentages, and indeed of manganese steel itself, are in some way due to actual combination between the iron and the manganese is, it will be understood, put forward in quite a tentative manner. In no other way does it seem possible at present to account for the practically complete suppression of the magnetism of the iron in such alloys. Explanations on the basis of allotropy of the iron are, as has been shown, quite inadequate; the allotropic changes of iron have in fact disappeared in these alloys. On the other hand, chemical combination of iron with other elements, it is known, does commonly affect its magnetic qualities, as also many of its other properties. For example, as is well known, while the oxide of iron Fe_3O_4 is magnetic, although not to the same degree as iron itself. Ferrous oxide, FeO, is non-magnetic. Further in the joint paper by the late Professor B. Hopkinson and the author on "The Magnetic Properties of Iron and its Alloys in Intense Fields," it was shown that iron carbide, Fe_3C, is about two-thirds as magnetic as pure iron. The double carbide, also of iron and manganese separated from manganese steel, was found by the author to be quite non-magnetic.

It is thus not unreasonable to suppose that manganese may effect a similar change in the magnetic qualities of iron even to the extent of their complete extinction, through the same medium of chemical combination. The suggestion

naturally does not in any way constitute an explanation in the true sense of the term. Granted the correctness of the suggestion, the explanation lies primarily in the answer to the still unsolved and fundamental question of why iron, nickel and cobalt exhibit ferromagnetism while other elements do not.

Without such knowledge it is obviously difficult, if not impossible, to indicate the actual mechanism by which such ferromagnetic properties become modified, when these elements are brought into association with others. The question thus becomes one essentially for the physicist or physical chemist. Professor Weiss, of Zurich, by his conception of the "magneton," which has formed the basis of much further work by others, did much towards probing the nature of ferromagnetism, and it is perhaps not too much to hope that we may, in the not too distant future, be led to a complete understanding and solution of this hitherto somewhat baffling problem.

In conclusion, the author desires to express his best thanks to Mr. S.A. Main, B.Sc., F.Inst.P., Mr. W. J. Todd, A.Inst.P., and Mr. A. Stevenson, who have with so much painstaking care assisted him in carrying through this laborious research.

Tensile Tests on Alloy Crystals. Part I.—*Solid Solution Alloys of Aluminium and Zinc.*

By C. F. Elam, Armourers' and Braziers' Research Fellow.

(Communicated by H. C. H. Carpenter, F.R.S.—Received March 29, 1927.)

[Plate 2.]

In the 'Proceedings of the Royal Society,' 1925,* some properties of crystals of an aluminium-zinc alloy were described and their method of extension and fracture compared with that of pure aluminium. This alloy contained 18·5 per cent. zinc and was just at the limit of solubility of zinc in aluminium. The crystals broke usually along a plane at, approximately, 45° to the axis and the position of this plane varied with regard to the crystal planes according to the orientation of the particular crystals. This type of fracture was so different to that of the pure metal that it appeared desirable to obtain alloys of intermediate compositions in order to study the effects of increasing amounts of zinc. Further, the small extension before fracture of the alloys containing higher percentages of zinc made it impossible to obtain reliable distortion measurements although measurements of the crystal axes, by means of X-rays, indicated that the distortion was probably of the same nature as aluminium. Through the courtesy of Mr. Murray Morrison, of the British Aluminium Company, a series of alloys were made in the form of 0·5 inch diameter bar. The number of the bar and the corresponding quantity of zinc is given in Table I. The bars will be referred to in future by the numbers in the first column.

Table I.

No.	Zinc.	No.	Zinc.
	per cent.		per cent.
5·0	5·45	15·0	16·10
7·5	7·11	17·0	17·59
10·0	10·60	18·0	18·50

The bars were cut into lengths, annealed, strained, and further annealed in a way similar to that which had been successful in growing large aluminium crystals. It was found much more difficult to grow crystals right through

* C. F. Elam, 'Roy. Soc. Proc.,' A, vol. 109, p. 143 (1925).

the cross-section than in aluminium and the results were very variable. This was to be explained partly by the fact that having stretched a bar in the machine the required amount (1 to 2 per cent. on 3 inches), on removing the load there was sometimes as much as 1 per cent. contraction. This applied mostly to alloys containing from 10 per cent. zinc upwards. The bars, therefore, tended to be understrained. If allowance was made for subsequent shrinking by over-stretching, equally variable results were obtained as the bars did not all shrink to the same amount even after carefully regulating the time of application of the load. The alloys as a whole are more sensitive to straining and to heat treatment than the pure metal, and small variations which appeared to make little difference to aluminium doubled or halved the crystal size. An extension of 1 to 2 per cent. was the most successful treatment with a previous and subsequent annealing temperature of 500–550° C. Several crystals occupying the whole cross section were usually obtained in one bar, and these were cut up and machined into small round test-pieces of 0·25 inches diameter, and of varying lengths according to the lengths of the crystal. The best crystals were put aside for machining square for distortion experiments. These will be described later.

The results of a number of tensile tests on these crystals showed that as the percentage of zinc increased, the amount of extension prior to fracture tended to decrease, while the breaking load increased, but there was considerable variation, especially in the elongation, between crystals of the same composition. In Table II are given the maximum breaking loads and elongations of a series of alloys in the normal multicrystalline state, and in Table III a few representative examples of the figures obtained from single crystals. These differences can be ascribed, partly to variation of crystal orientation, and partly to the possibility of a slightly varying composition from bar to bar.

All these alloys show a well-marked yield-point, particularly in the higher percentages of zinc, and it was more noticeable in single crystal test-pieces than in the normal material.

All the crystals tended to become elliptical in cross-section. As has been stated already, the alloys containing 18·5 per cent. zinc break along a nearly straight cleavage plane, while pure aluminium in common with other similar metals, such as copper and silver, pull down to a double wedge at the point of fracture. This has been shown* to be due to the fact that the distortion of the crystal takes place by means of slip on the octahedral plane subjected

* Taylor and Elam, ' Roy. Soc. Proc.,' A, vol. 102 (1923).

Table II.—Normal annealed Test-pieces.

No.	Maximum Load.	Elongation per cent. on 2 inches.
	Tons per square inch.	
Pure aluminium	4·82	44·0
5 { 1	6·28	34·0
5 { 2	6·29	33·5
10 { 1	10·00	31·0
10 { 2	9·79	28·0
15 { 1	15·41	24·5
15 { 2	15·25	25·0
17 { 1	15·45	19·0
17 { 2	15·35	19·0
18 { 1	19·02	12·0
18 { 2	19·12	13·0

Table III.—Single crystals.

No.		Maximum load.	Elongation per cent. on 1 inch.	No.	Maximum load.	Elongation per cent. on 1 inch.
		Tons per square inch.			Tons per square inch.	
5.	1	4·46	58	15. 1	15·81	26
	23	4·835	31	3	15·01	17
	61	5·13	58	24	14·50	48
7·5.	5	5·84	33	18. 20	24·95	0·5
	2	6·07	34	21	20·55	8·0
10.	1	11·65	62	22	19·20	7·0
	3	9·32	72			
	55	11·90	20			
	49	11·49	31			
	43	9·56	57			
	41	11·09	54			

to the greatest shear stress and towards the pole of the {110} plane nearest the direction of greatest slope, until by the rotation of the crystal, a second octahedral plane makes an equal angle with the axis as the first, when slip occurs on both planes.

Alloys containing 5 to 7·5 per cent. zinc usually show such a wedge-shaped fracture, while alloys containing 10 per cent. zinc and upwards showed a straight cleavage fracture (fig. 1, Plate 2) similar to that described in the previous

paper. Certain crystals containing 5 to 7·5 per cent. zinc showed an intermediate structure in which the "double" wedge was replaced by a single wedge asymmetric to the axis of the specimen. A photograph illustrating this structure is shown in fig. 2, Plate 2. The explanation is thought to be due to the fact that as the zinc content increases, the possible amount of extension before fracture decreases so that fracture may occur before the position for slip on two planes is reached, and in that case a single plane fracture results. This is confirmed by the results of distortion measurements in most cases. Occasionally a fracture such as is shown in fig. 3, Plate 2, is obtained, which closely resembles the double wedge fracture, but clearly consists of two planes. In such specimens, when the crystal axes have been determined, they have been found to be near the position for double slipping. The intermediate types of fracture in some of the alloys containing less zinc appear to be due to their greater ductility, and to a rather different type of distortion at the moment of fracture when the metal seems to flow over the plane of fracture. The relation existing between the plane of fracture and the crystal structure can best be discussed after describing the results of distortion measurements.

Fig. 4, Plate 2, shows examples of all the alloy crystals.

Distortion and X-Ray Measurements.

A few of the best crystals were machined square for distortion measurements. A complete set of measurements was made on one crystal containing 10 per cent. zinc, and two containing 15 per cent. zinc. These were machined 0·25 inch square and were 4 to 6 inches long. The methods of marking and measuring were the same as those previously adopted and described,* and calculations were made of the cone of unextended directions for different stages of the extension. The 10 per cent. zinc alloy crystal extended 63 per cent. before fracture so that a number of measurements could be made. Of the 15 per cent. zinc alloy crystals, two series of measurements were made of the first, while only one was made of the second as this broke after 15 per cent. extension. The results of both measurements and calculations are given in the following tables. (Table IV.)

The notation is the same as that used in the previous paper, but is briefly described here. The four faces of the specimens were marked I, II, III and IV, so that Face II was 90° in an anti-clockwise direction from Face I, when

* Taylor and Elam, 'Roy. Soc. Proc.,' A, vol. 112, p. 337 (1926).

the specimen was held vertical. The following letters represent particular measurements or values :—

ε = ratio of extended to initial length.

f = ratio of width of Face I, after extension, to its original width.

g = ratio of width of Face IV after extension to its original width.

λ_0 = Angle between Face I and Face IV.

β_0 = Angle between scratches on Face I and vertical axis of specimen.

γ_0 = Angle between scratches on Face IV and vertical axis of specimen.

λ_1 β_1 and γ_1 are the corresponding values after extension.

θ = Angle which a given direction makes with the vertical axis.

ϕ = Angle which the projection of the direction on a plane perpendicular to the axis makes with the reference plane. ϕ is measured anti-clockwise.

Table IV.

Crystal I. 10 per cent. zinc. Reference Face I.

$\epsilon = 1{\cdot}114$, $f = 1{\cdot}001$, $g = 0{\cdot}8975$, $\lambda_0 = 90°$, $\lambda_1 = 88{\cdot}2°$
$\beta_0 = 90° 10'$, $\beta_1 = 88° 22'$, $\gamma_0 = 89° 50'$, $\gamma_1 = 89°$

Unstretched cone.

Unstrained position.

$$(0{\cdot}003 \cos^2\phi - 0{\cdot}1987 \sin^2\phi + 0{\cdot}0581 \cos\phi \sin\phi)\tan^2\theta + (0{\cdot}0709 \cos\phi + 0{\cdot}0278 \sin\phi)\tan\theta + 0{\cdot}241 = 0.$$

$\epsilon = 1{\cdot}204$, $f = 1{\cdot}000$, $g = 0{\cdot}831$, $\lambda_0 = 90°$, $\lambda_1 = 86{\cdot}5°$
$\beta_0 = 90° 10'$, $\beta_1 = 86° 30'$, $\gamma_0 = 89° 50'$ $\gamma_1 = 89° 40'$

Unstretched cone.

Unstrained position.

$$(0{\cdot}0031 \cos^2\phi - 0{\cdot}3085 \sin^2\phi + 0{\cdot}1029 \cos\phi \sin\phi)\tan^2\theta + (0{\cdot}1542 \cos\phi + 0{\cdot}0033 \sin\phi)\tan\theta + 0{\cdot}452 = 0.$$

Strained position.

$$(0{\cdot}00289 \cos^2\phi + 0{\cdot}4527 \sin^2\phi - 0{\cdot}1227 \cos\phi \sin\phi)\tan^2\theta - (0{\cdot}0892 \cos\phi + 0{\cdot}0044 \sin\phi)\tan\theta - 0{\cdot}310 = 0.$$

$\epsilon = 1{\cdot}408$, $f = 0{\cdot}9975$, $g = 0{\cdot}7195$, $\lambda_0 = 90°$, $\lambda_1 = 85{\cdot}2°$
$\beta_0 = 90° 10'$, $\beta_1 = 83° 30'$, $\gamma_0 = 89° 50'$, $\gamma_1 = 95° 40'$

Unstretched cone.

Unstrained position.

$$(0{\cdot}0078 \cos^2\phi - 0{\cdot}4777 \sin^2\phi + 0{\cdot}1141 \cos\phi \sin\phi)\tan^2\theta + (0{\cdot}3305 \cos\phi - 0{\cdot}2117 \sin\phi)\tan\theta + 0{\cdot}975 = 0$$

$\epsilon = 1{\cdot}637$ to $\epsilon = 1{\cdot}408$.
$\epsilon = 1{\cdot}162$, $f = 1{\cdot}000$, $g = 0{\cdot}8610$, $\lambda_0 = 85{\cdot}2°$, $\lambda_1 = 85°$
$\beta_0 = 83\ °30'$, $\beta_1 = 78° 10'$, $\gamma_0 = 95° 40'$, $\gamma_1 = 104° 50'$.

Unstretched cone.

Strained position.

$$(0{\cdot}0043 \cos^2\phi + 0{\cdot}3722 \sin^2\phi - 0{\cdot}0396 \cos\phi \sin\phi)\tan^2\theta - (0{\cdot}1138 \cos\phi + 0{\cdot}2315 \sin\phi)\tan\theta - 0{\cdot}260 = 0.$$

Table IV—(continued).

Crystal I. 15 per cent. zinc.* Reference Face IV.

$\epsilon = 1{\cdot}097$, $f = 0{\cdot}9950$, $g = 0{\cdot}9175$, $\lambda_0 = 89{\cdot}7°$, $\lambda_1 = 86{\cdot}1°$
$\beta_0 = 90° 10'$, $\beta_1 = 93° 40'$, $\gamma_0 = 90° 20'$, $\gamma_1 = 88° 30'$

Unstretched cone.
Unstrained position.

$$(0{\cdot}0024 \cos^2\phi + 0{\cdot}1589 \sin^2\phi - 0{\cdot}1214 \cos\phi \sin\phi) \tan^2\theta + (0{\cdot}133 \cos\phi - 0{\cdot}0666 \sin\phi) \tan\theta - 0{\cdot}201 = 0.$$

Strained position.

$$(0{\cdot}0131 \cos^2\phi + 0{\cdot}1963 \sin^2\phi - 0{\cdot}1326 \cos\phi \sin\phi) \tan^2\theta + (0{\cdot}1012 \cos\phi - 0{\cdot}0627 \sin\phi) \tan\theta - 0{\cdot}167 = 0.$$

$\epsilon = 1{\cdot}222$, $f = 0{\cdot}9975$, $g = 0{\cdot}8590$, $\lambda_0 = 89{\cdot}7$, $\lambda_1 = 81{\cdot}8$.
$\beta_0 = 90° 10'$, $\beta_1 = 97° 30'$, $\gamma_0 = 90° 20'$, $\gamma_1 = 89°$.

Unstretched cone.
Unstrained position.

$$(0{\cdot}0112 \cos^2\phi - 0{\cdot}2657 \sin^2\phi + 0{\cdot}2338 \cos\phi \sin\phi) \tan^2\theta - (0{\cdot}3115 \cos\phi + 0{\cdot}0550 \sin\phi) \tan\theta + 0{\cdot}493 = 0.$$

Strained position.

$$(0{\cdot}0249 \cos^2\phi + 0{\cdot}3889 \sin^2\phi - 0{\cdot}2983 \cos\phi \sin\phi) \tan^2\theta + (0{\cdot}2585 \cos\phi - 0{\cdot}0258 \sin\phi) \tan\theta - 0{\cdot}317 = 0.$$

Crystal II. Reference Face I.

$\epsilon = 1{\cdot}145$, $f = 0{\cdot}8770$, $g = 1{\cdot}000$, $\lambda_0 = 90°$, $\lambda_1 = 89°$
$\beta_0 = 88° 30'$, $\beta_1 = 91° 20'$, $\gamma_0 = 89° 40'$, $\gamma_1 = 87° 40'$.

Unstretched cone.
Unstrained position.

$$(0{\cdot}2285 \cos^2\phi - 0{\cdot}0010 \sin^2\phi - 0{\cdot}0238 \cos\phi \sin\phi) \tan^2\theta + (0{\cdot}1153 \cos\phi - 0{\cdot}0781 \sin\phi) \tan\theta - 0{\cdot}311 = 0.$$

Strained position.

$$(0{\cdot}3054 \cos^2\phi + 0{\cdot}0013 \sin^2\phi - 0{\cdot}0487 \cos\phi \sin\phi) \tan^2\theta + (0{\cdot}0861 \cos\phi - 0{\cdot}0535 \sin\phi) \tan\theta - 0{\cdot}237 = 0.$$

There are two positions for the unstretched cone; one in the unstrained position, and one in the strained. In some cases, only one was calculated as this is sufficient to show if the crystal is slipping on one plane or two, and if it agrees with the X-ray measurements. With one exception, the distortion was calculated from the measurements when $\varepsilon = 1{\cdot}0$, but for the last stage of extension of the 10 per cent. alloy calculations were made between measurements at $\varepsilon = 1{\cdot}408$ and $\varepsilon = 1{\cdot}637$ in order to see if the last stage of slip was different.

In Table V are given the two calculated values of θ for the cone of unextended directions in the unstrained position for extensions up to $\varepsilon = 1{\cdot}408$ in the case of the 10 per cent. zinc alloy, and up to the breaking point in the case of one 15 per cent. zinc alloy. The values are almost constant for one plane, indicating that this is the slip-plane. In Table VI the position of the pole of

* In this particular crystal measurements were taken from Faces III and IV, so that λ_0 = angle between Faces III and IV, etc.

the slip-plane calculated from distortion measurements is compared with that of the nearest {111} plane, and the direction of slip* with the pole of the nearest {110} plane. The agreement is not so good in the second 15 per cent. crystal, which showed a tendency to twist, specially near the fracture. The distortion was not on the whole so regular as in aluminium, due in part to the presence of small unabsorbed crystals in the main crystal. The results from both distortion and X ray measurements, however, indicate that the alloy crystals deform in the same way as pure aluminium, viz., by slip on an octahedral {111} plane in a direction represented by the normal to a {110} plane. This is further confirmed by observation of slip-bands. In every case only one set of slip-bands was observed, which agreed with the plane of slip calculated from distortion measurements. There is also evidence that the 10 per cent. alloy crystal slipped on a second plane for about the last 10 per cent. of its extension. Distortion measurements of the last 10 per cent. extension do

Table V.—10 per cent. zinc crystal.

Cone of unextended directions in " unstrained " position.

	$\epsilon = 1{\cdot}114$.		$\epsilon = 1{\cdot}204$.		$\epsilon = 1{\cdot}408$.	
ϕ	θ_1	θ_2	θ_1	θ_2	θ_1	θ_2
90°	49° 30′	135° 0′	50° 30′	129° 40′	51′ 0′	120° 50′
60°	58° 55′	151° 40′	60° 0′	127° 0′	62° 0′	118° 55′
30°	79° 25′	116° 45′	82° 0′	123° 30′	79° 50′	110° 0′
0°	91° 20′	85° 10′	107° 20′	91° 10′	107° 35′	91° 30′
−30°	123° 20′	65° 20′	124° 30′	68° 30′	123° 15′	75° 40′
−60°	131° 20′	50° 30′	131° 0′	54° 50′	130° 20′	64° 0′

15 per cent. zinc crystal.

	$\epsilon = 1{\cdot}097$.		$\epsilon = 1{\cdot}222$.	
ϕ	θ_1	θ_2	θ_1	θ_2
90°	53° 30′	137° 0′	54° 30′	128° 18′
60°	57° 20′	130° 55′	60° 30′	109° 10′
30°	—	—	—	—
0°	91° 0′	54° 30′	86° 40′	83° 50′
−30°	112° 10′	41° 10′	108° 55′	46° 35′
−60°	123° 45′	37° 30′	121° 20′	44° 20′

* The direction of slip is at right angles to the line of intersection of the two unstretched planes.

not bring this out to any marked extent, probably because the amount of slip on this plane was small, but the cone does not consist exactly of two planes. Compare figs. 1 and 2.

It has not been thought necessary to give a complete set of stereographic diagrams showing the position of the crystal planes, slip-plane, etc., as these are similar to others that have been published in the case of aluminium. Fig. 1

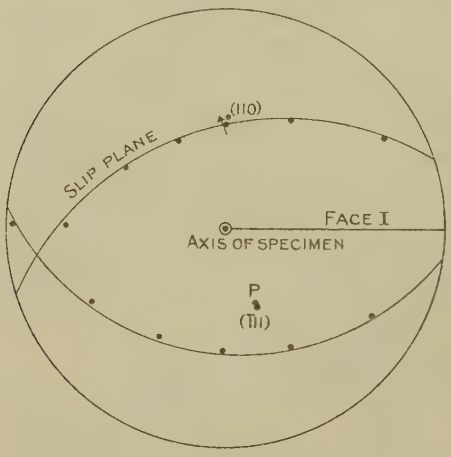


Fig. 1.—10 per cent. zinc.
$\epsilon = 1{\cdot}408$. Unstrained position.

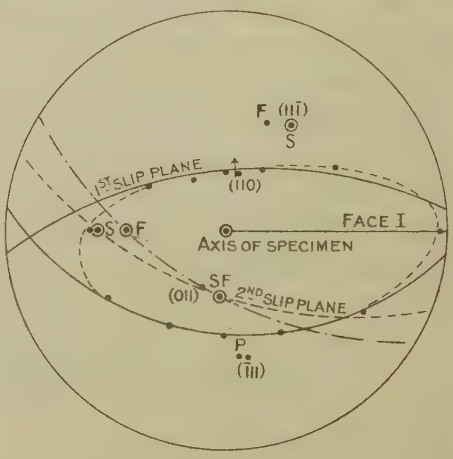


Fig. 2.—10 per cent. zinc.
$\epsilon = 1{\cdot}637$ to $1{\cdot}408$. Strained position.

shows the unstrained position of the cone of unextended directions when $\varepsilon = 1{\cdot}408$ for the 10 per cent. alloy, and fig. 2 the "strained" position when $\varepsilon = 1{\cdot}637$, representing the last stage of deformation from $\varepsilon = 1{\cdot}408$. The positions of slip-bands and fracture are here shown. Fig. 3 shows a similar

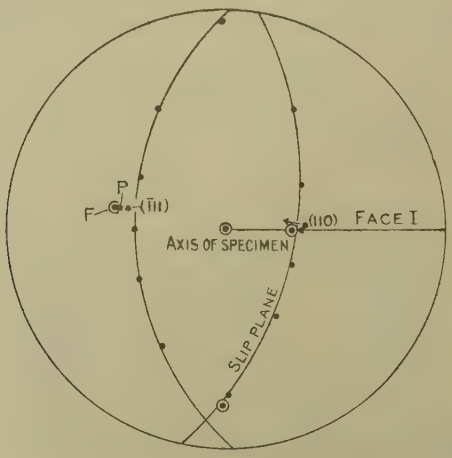


Fig. 3.—15 per cent. zinc. $\epsilon = 1{\cdot}145$. Strained position.

figure in the case of the second 15 per cent. zinc alloy, and also shows the position of slip-bands and fracture.

The pole of the slip-plane determined by distortion measurements is called P, that representing the slip-plane determined from the measurements of traces of slip-bands on the surface (marked ⊙ in diagrams) is called S, that of fracture F. The direction of slip is marked by an arrow. The indices used in referring to the planes are chosen arbitrarily, but are used consistently throughout the present paper to facilitate comparison between the several diagrams.

In figs. 4 and 5 are shown the positions of the axes of the test-pieces relative to the crystal of the three specimens investigated. The first and last positions are given, but not all the intermediate positions as these are not necessary in order to show the general movement of the axes. As has been shown previously* the axis of the specimen should move along a great circle joining its position before extension to the pole of the (110) plane towards which slip takes place. Where this great circle cuts that passing through the poles of the (010) and (111) planes, the axis is equally inclined to two octahedral planes ($\bar{1}11$) and ($11\bar{1}$), the former of which should be the first slip plane, and slip should occur on both planes simultaneously. The 10 per cent. alloy should have reached this point at an extension of about 50 per cent. When $\varepsilon = 1{\cdot}55$ the axis was in this position (*see* fig. 5) and remained almost constant up to the breaking point at $\varepsilon = 1{\cdot}637$. This is in agreement with the distortion measurements during the last stage of the extensions already referred to, with the appearance of a second set of slip-bands which coincide with the trace of the second octahedral plane, and with the fact that the plane of fracture agreed more nearly with this plane than with the first octahedral plane. It is also in agreement with what has previously been observed in aluminium, copper, silver and gold. All these points are made clearer by referring to the stereographic diagram, fig. 2.

The positions of the planes of the fracture have been given in Table VI. In many specimens the plane of fracture agreed very closely with the slip-plane, but in many the two made a small angle with each other, usually of from 10° to 20°. This is noticeable in some of the photographs, figs. 1, 2, 3 and 6, Plate 2, and was observed in the previous work on the 18·5 per cent. zinc crystals.

Fig. 4 shows that the two crystals containing 15 per cent. zinc, broke long before they reached the position for double slipping. It also shows that the direction of slip was not quite in the direction demanded by theory, although

* G. I. Taylor and C. F. Elam, 'Roy. Soc. Proc.,' A, vol. 108 (1925).

distortion measurements agree very closely with slip towards the pole of the (110) plane (*see* fig. 3). The error is probably in the first X-ray measurements of the crystal.

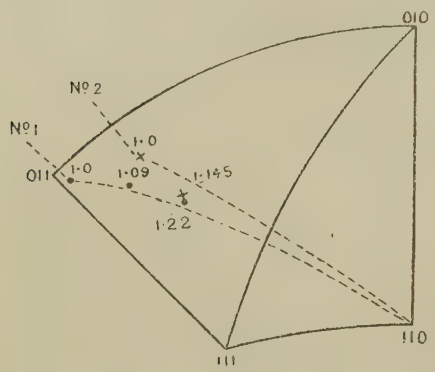


Fig. 4.—15 per cent. zinc.

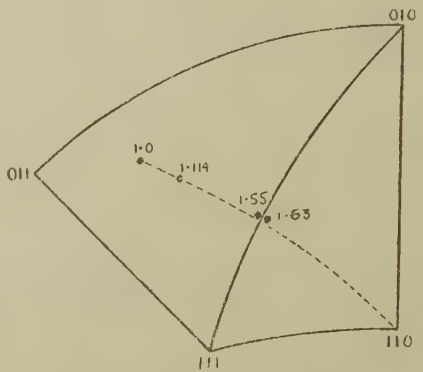


Fig. 5.—10 per cent. zinc.

Table VI.

10 per cent. zinc crystal I. Reference Face I.

Slip-plane.		Direction of slip.	
X-ray measurements.	Distortion measurements.	X-ray measurements.	Distortion measurements.
$\epsilon = 1{\cdot}114$. Unstrained			
$\theta_{\bar{1}11} = 43° 20'$	$\theta = 43° 0'$	$\theta_{110} = 52° 20'$	$\theta = 49° 0'$
$\phi = 291° 20'$	$\phi = 288° 30'$	$\phi = 91° 0'$	$\phi = 90° 0'$
$\epsilon = 1{\cdot}204$. Unstrained			
$\theta_{\bar{1}11} = 43° 20'$	$\theta = 41° 20'$	$\theta_{110} = 52° 20'$	$\theta = 51° 10'$
$\phi = 291° 20'$	$\phi = 290° 10'$	$\phi = 91° 0'$	$\phi = 90° 0'$
Strained			
$\theta_{\bar{1}11} = 51° 0'$	$\theta = 52° 0'$	$\theta_{110} = 49° 10'$	$\theta = 49° 40'$
$\phi = 281° 20'$	$\phi = 284° 20'$	$\phi = 92° 30'$	$\phi = 87° 0'$
$\epsilon = 1{\cdot}408$. Unstrained			
$\theta_{\bar{1}11} = 43° 20'$	$\theta = 41° 30'$	$\theta_{110} = 52° 20'$	$\theta = 50° 40'$
$\phi = 291° 20'$	$\phi = 290° 30'$	$\phi = 91° 0'$	$\phi = 91° 0'$
$\epsilon = 1{\cdot}637$ Strained			
$\theta_{\bar{1}11} = 61° 10'$	$\theta = 61° 0'$	$\theta_{110} = 30° 0'$	$\theta = 30° 40'$
$\phi = 281° 0'$	$\phi = 287° 0'$	$\phi = 78° 0'$	$\phi = 80° 0'$
Unstrained			
$\theta_{111} = 43° 20'$	$\theta = 42° 0'$	$\theta_{110} = 52° 20'$	$\theta = 52° 10'$
$\phi = 291° 20'$	$\phi = 291° 40'$	$\phi = 91° 0'$	$\phi = 90° 40'$
Fracture	$\theta = 56° 0'$	Slip-bands	$\theta = 59° 30'$
	$\phi = 69° 0'$	at fracture	$\phi = 59° 10'$

Table VI—(continued).

15 per cent. zinc crystal I. Reference Face IV.

Slip-plane.		Direction of Slip.	
X-ray measurements.	Distortion measurements.	X-ray measurements.	Distortion measurements.
$\epsilon = 1{\cdot}097$ Unstrained			
$\theta_{\bar{1}11} = 35° 30'$	$\theta = 37° 0'$	$\theta_{110} = 59° 0'$	$\theta = 56° 30'$
$\phi = 274° 0'$	$\phi = 271° 0'$	$\phi = 122° 0'$	$\phi = 120° 30'$
Strained			
$\theta_{\bar{1}11} = 41° 30'$	$\theta = 41° 30'$	$\theta_{110} = 50° 40'$	$\theta = 50° 20'$
$\phi = 276° 10'$	$\phi = 279° 0'$	$\phi = 121° 30'$	$\phi = 119° 10'$
$\epsilon = 1{\cdot}222$ Unstrained			
$\theta_{\bar{1}11} = 35° 30'$	$\theta = 34° 10'$	$\theta_{110} = 59° 0'$	$\theta = 60° 40'$
$\phi = 274° 0'$	$\phi = 270° 40'$	$\phi = 122° 0'$	$\phi = 125° 0'$
Strained			
$\bar{1}11 = 48° 30'$	$\theta = 47° 0'$	$\theta_{110} = 55° 0'$	$\theta = 59° 0'$
$\phi = 278° 0'$	$\phi = 270° 0'$	$\phi = 124° 30'$	$\phi = 128° 0'$
Fracture	$\theta = 49° 0'$	Slip-bands	$\theta = 48° 30'$
	$\phi = 256° 0'$	at fracture	$\phi = 263° 0'$

15 per cent. zinc crystal II. Reference Face I.

Slip-plane.		Direction of slip.	
X-ray measurements.	Distortion measurements.	X-ray measurements.	Distortion measurements.
$\epsilon = 1{\cdot}145$ Unstrained			
$\theta_{\bar{1}11} = 42° 0'$	$\theta = 48° 0'$	$\theta_{110} = 49° 20'$	$\theta = 43° 0'$
$\phi = 165° 20'$	$\phi = 169° 30'$	$\phi = 7° 30'$	$\phi = 4° 50'$
Strained			
$\theta_{\bar{1}11} = 49° 05'$	$\theta = 52° 0'$	$\theta_{110} = 39° 30'$	$\theta = 36° 30'$
$\phi = 169° 40'$	$\phi = 170° 0'$	$\phi = 2° 30'$	$\phi = 1° 30'$
Fracture	$\theta = 54° 0'$	Slip-bands	$\theta = 54° 0'$
	$\phi = 170° 0'$	at fracture	$\phi = 170° 0'$

Fig. 5, Plate 2, shows the beginning of the fracture of the 10 per cent. alloy. A second set of slip-bands suddenly appeared which were only visible on one pair of faces so that the plane could not be measured. The inclination of this "fault" agreed with the general direction of the slip plane, but when fracture was complete, the two were seen to be quite distinct, fig. 6, Plate 2. It is as if fracture began by failure on the slip-plane, but that some distortion, probably on one or more other planes, led to the final rupture. The type of fracture already referred to in the 5 to 7·5 per cent. alloys, and illustrated in

fig. 2, Plate 2, may also be caused by some such distortion. This particular crystal broke before it had reached the position for double slipping so that a second plane was not involved in the formation of a double wedge.

Fig. 7, Plate 2, is a photograph of the 15 per cent. zinc alloy after fracture.

The shear stress (S*) on the slip plane and the amount of shear (s) were calculated for some of these crystals. The positions of the crystal axes were determined before testing and the slip-plane deduced from previous knowledge of the behaviour of these crystals. From this the angle which the normal to the slip-plane made with the test-piece axis (θ) and the angle between the direction of slip and the direction of greatest slope (η) were calculated. The load and cross-section were measured in the usual way. The amount of shear is calculated from the formula :—

$$\varepsilon^2 = 1 + 2s \cos \eta_0 \sin\theta_0 \cos \theta_0 + s^2 \cos^2 \theta_0$$

where s is the displacement of a slip-plane relative to a parallel plane at unit distance from it.† By using this method of plotting shear-stress against amount of shear, variations which can be attributed to differences of orientation in the crystals compared can be eliminated.

The results are given in Table VII and shown diagrammatically in fig. 6. These show clearly the effect of zinc in increasing the resistance to shear. Some figures for pure aluminium are given for comparison. The maximum increase in hardness occurs in the early stages of deformation, but the shear

Table VII.

Pure aluminium.

$\theta_{\bar{1}11} = 41° 0'$, $\eta = 12° 0'$ } Slip-plane.

ϵ	S	s
	lbs. per sq. inch.	
1·053	2,180	0·1241
1·110	2,930	0·2088
1·161	3,380	0·3045
1·200	3,740	0·3700
1·304	4,140	0·5420
1·404	4,480	0·7065
1·623	5,030	1·043
1·785	5,570	1·282

* *Loc. cit.*, ' Roy. Soc. Proc.,' A, vol. 112 (1923).

† I am indebted to Prof. G. I. Taylor, F.R.S., for this method of calculating the amount of shear.

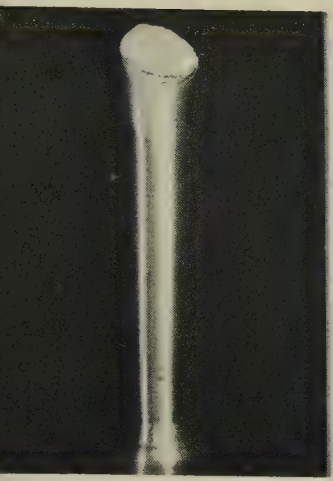


FIG. 1.—15 per cent. Zn.

FIG. 2.—5 per cent. Zn.

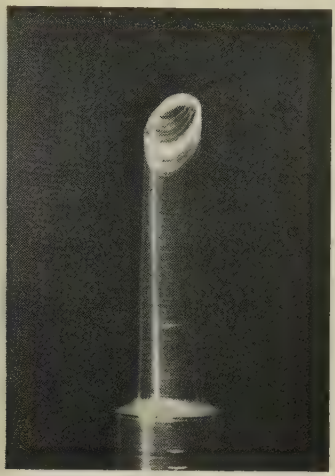
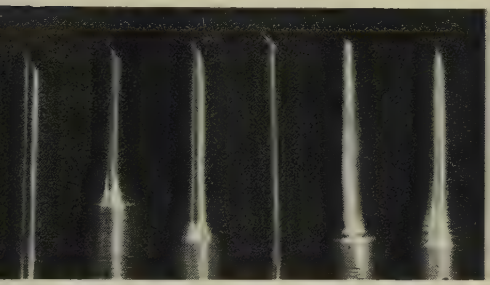


FIG. 3.—10 per cent. Zn.



18·5 17 15 10 7·5 5 per cent. Zn.

FIG. 4.

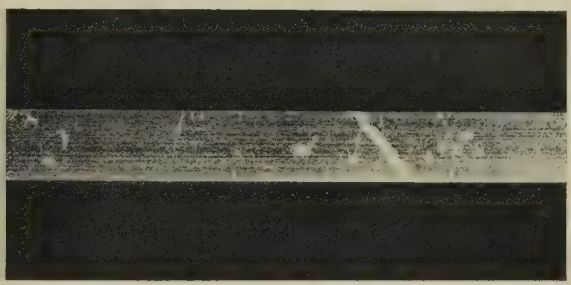


FIG. 5.—Beginning of Fracture in 10 per cent. Zn. Alloy Crystal.

FIG. 6.—10 per cent. Zn. Alloy broken.

FIG. 7.—15 per cent. Zn. Alloy broken.

(*Facing p.* 144.)

Table VII—(continued).

Aluminium-zinc.—5 per cent. zinc No. 5—61.

$\theta_{111} = 44° 30'$
$\eta = 30° 0'$ } Slip-plane.

ϵ.	S.	s.
	lbs. per sq. inch.	
1·05	2,693	0·1189
1·11	3,910	0·2340
1·17	4,710	0·3501
1·25	5,585	0·5010
1·37	6,295	0·7170
1·45	6,545	0·8500
1·58	7,010	1·066

Aluminium-zinc.—10 per cent. zinc No. I.

$\theta_{\bar{1}11} = 43° 0'$
$\eta = 12° 0'$ } Slip-plane.

ϵ.	S.	s.
	lbs. per sq. inch.	
1·114	10,850	0·2025
1·204	12,380	0·3725
1·305	12,920	0·5245
1·408	13,480	0·7055
1·50	13,920	0·8530
1·637	15,230	1·062

No. II.

$\theta_{\bar{1}11} = 45° 20'$
$\eta = 27° 30'$ } Slip-plane.

ϵ.	S.	s.
	lbs. per sq. inch.	
1·05	7,010	0·1082
1·11	8,710	0·2332
1·15	9,790	0·3107
1·22	10,610	0·4435
1·32	11,330	0·6245
1·45	11,830	0·8450
1·53	12,520	0·9790

Table VII—(continued).

Aluminium-zinc.—15 per cent. zinc No. I.

$\theta_{\bar{1}11} = 37° 0'$
$\eta = 23° 30'$ } Slip-plane.

ϵ.	S.	s.
	lbs. per sq. inch.	
1·097	14,380	0·1987
1·222	16,770	0·4315

No. II.

$\theta_{\bar{1}11} = 42° 0'$
$\eta = 17° 30'$ } Slip-plane.

ϵ.	S.	s.
	lbs. per sq. inch.	
1·145	17,900	0·2788

No. III.

$\theta_{\bar{1}11} = 42° 30'$
$\eta = 2°$ } Slip-plane.

ϵ.	S.	s.
	lbs. per sq. inch.	
1·05	12,300	0·0548
1·08	14,000	0·1539
1·15	16,500	0·2935
1·24	17,920	0·4350
1·35	19,130	0·6185
1·48	20,380	0·8295

Aluminium-zinc.—18 per cent. zinc.

$\theta_{\bar{1}11} = 48° 20'$
$\eta = 20° 30'$ } Slip-plane.

ϵ.	S.	s.
	lbs. per sq. inch.	
1·014	37,080	0·0320
1·028	40,850	
1·056	45,050	0·1157
1·082	46,550	0·1709
1·098	47,050	0·1981
1·110	47,100	0·2219

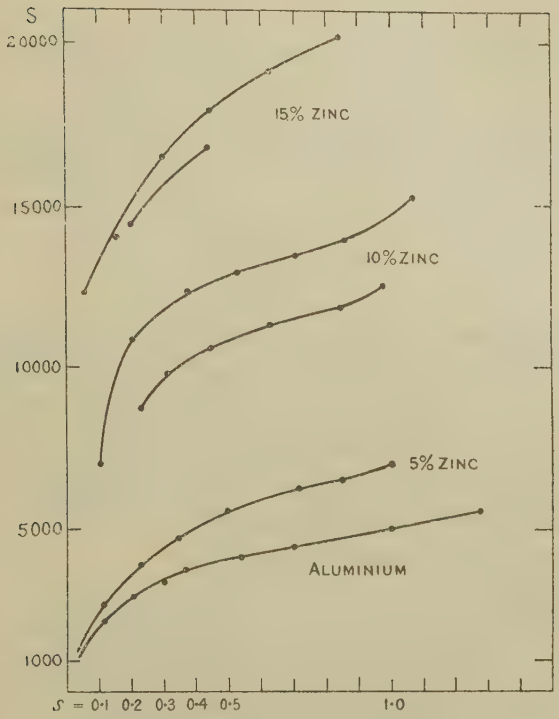


FIG. 6.— S = Shearing stress measured in pounds per square inch.
s = Amount of shear.

stress increases right up to the breaking point. There is a tendency for the final value to be rather exceptionally high. This has been observed in a number of test-pieces of other metals, and may be due to a difference of deformation at or near the breaking point.

The curves for two 10 per cent. zinc crystals differ considerably in value though not in shape, and attention has already been drawn to the variable results obtained amongst these alloys.

Tensile Tests on Alloy Crystals. Part II.—*Solid Solution Alloys of Copper and Zinc.*

By C. F. Elam, Armourers' and Braziers' Research Fellow.

(Communicated by H. C. H. Carpenter, F.R.S.—Received March 29, 1927.)

[Plates 3 and 4.]

A previous paper* has described the method of making rods of metal in the form of one crystal by slowly lowering graphite tubes containing a rod of the required metal through a furnace, so that cooling and solidification begins from the bottom of the tube. Copper, silver and gold crystals were made by this method and showed the same properties on distortion as aluminium crystals made by another method. The present paper describes some tensile tests carried out on brass crystals made in this way.

Some rods 0·25 inch in diameter of a brass containing 70 per cent. copper and 30 per cent. zinc were kindly given to the author by Mr. Leonard Sumner, of the Broughton Copper Company, Manchester. Lengths of 9 inches exactly fitted into graphite tubes and these were lowered through a platinum wound furnace which was maintained at 1,150° C. Nitrogen was passed slowly through the tube to prevent oxidation. The whole process was carried out as quickly as possible in order to reduce the loss of zinc. In spite of this a considerable amount of zinc distilled off and collected at the end of the tube and on the nickel-chrome wire by which the graphite tube was suspended. This made the wire very brittle and if the temperature were too high or the wire too long in the furnace, it broke away. A number of crystals were successfully prepared and the copper was estimated in samples from each end of the rod. As it was desired to preserve the part upon which distortion measurements had been made it was only possible to cut off a small piece at each end of the crystal where it had been held in the testing machine, and in the case of the top of the ingot this was not always sound. Only one analysis was made of Crystal I as the portion measured was comparatively small, near the centre of the rod, and the sample for analysis was taken as close to this as possible. As was to be expected, very variable results were obtained, according to the time taken in melting and lowering the crystal, and the temperature of the furnace. The differences obtaining between each end were also variable,

* C. F. Elam, 'Roy. Soc. Proc.,' A, vol. 112, p. 289 (1926).

ranging from 0·51 per cent. to 3·28 per cent. in the amount of copper. It will also be seen that the amount of zinc lost, judged by the copper content, varies from about 1 per cent. to 10 per cent. Some of the crystals had a bright-yellow colour at the top, and it is probable that this was due to the increased amount of zinc which distilled out of the rest of the metal and re-condensed as the tube cooled.

Table I.

Crystal No.	Copper, per cent.	
	Top.	Bottom.
I.	70·90	
II.	74·47	77·75
III.	72·62	74·68
IV.	75·25	76·10
V.	80·20	80·71

It is probable that the portion of the crystals measured is in closer agreement with the analysis of the bottom end than the top.

Westgren and Phragmen* have made a careful study of the alloys of copper and zinc, and find that zinc atoms replace copper atoms in the face-centred cubic lattice over the range of α solid solutions. This includes alloys up to 36 per cent. zinc. They give the side of the unit cube for an alloy containing 32·2 per cent. zinc as 3·688 Å, compared with 3·610 Å for pure copper, *i.e.*, an increase of about 2 per cent. The crystals used in the present research had rather less zinc than this but the differences in lattice dimensions are very small.

In spite of the low rate of cooling, there was a slight appearance of choring. In order to remove this, the crystals were heated at 750° to 800° C. for four days. At the end of that time polishing and etching did not reveal any signs of non-homogeneity. The crystals were not sound, however, and small blow-holes occurred in certain directions agreeing, roughly, with the original planes of choring. In order to prepare for distortion measurements they were machined square and all four sides carefully polished and etched until all signs of tool marks were removed. The final size was 3 to 4 mms. square and about 10 cms. long. They were then marked and measured in the way already described.†

The results of measurements and calculations are given in Table II.

* Westgren and Phragmen, ' Phil. Mag.,' vol. 1, July, 1925.

† G. I. Taylor and C. F. Elam, ' Roy. Soc. Proc.,' A, vol. 112 (1926).

Table II.

α Brass I.—Reference Face I.

1st extension.

$\epsilon = 1{\cdot}175$, $f = 1{\cdot}000$, $g = 0{\cdot}8590$, $\lambda_0 = 89{\cdot}5°$, $\lambda_1 = 88{\cdot}5°$,
$\beta_0 = 90°10'$, $\beta_1 = 86°\ 30'$, $\gamma_0 = 90°$, $\gamma_1 = 99°\ 30'$.

Unstretched cone.
Unstrained position.

$$(0{\cdot}0042\ \cos^2\phi - 0{\cdot}2356\ \sin^2\phi + 0{\cdot}0263\ \cos\phi\ \sin\phi)\ \tan^2\theta + (0{\cdot}1517\ \cos\phi - 0{\cdot}3395\ \sin\phi)\ \tan\theta + 0{\cdot}380 = 0.$$

Strained position.

$$(0{\cdot}0030\ \cos^2\phi + 0{\cdot}3708\ \sin^2\phi - 0{\cdot}0478\ \cos\phi\sin\phi)\ \tan^2\theta - (0{\cdot}0934\ \cos\phi - 0{\cdot}2442\ \sin\phi)\ \tan\theta - 0{\cdot}278 = 0.$$

2nd extension.

$\epsilon = 1{\cdot}272$, $f = 1{\cdot}000$, $g = 0{\cdot}7965$, $\lambda_0 = 89{\cdot}5°$, $\lambda_1 = 88{\cdot}5°$,
$\beta_0 = 90°$, $\beta_1 = 83°\ 40'$, $\gamma_0 = 90°$, $\gamma_1 = 106°\ 10'$.

Unstretched cone.
Unstrained position.

$$(0{\cdot}0115\ \cos^2\phi - 0{\cdot}3118\ \sin^2\phi - 0{\cdot}0279\ \cos\ \phi\sin\phi)\ \tan^2\theta + (0{\cdot}2739\ \cos\phi - 0{\cdot}5895\ \sin\phi)\ \tan\theta - 0{\cdot}620 = 0.$$

Strained position.

$$(0{\cdot}0081\ \cos^2\phi + 0{\cdot}6321\ \sin^2\phi - 0{\cdot}0772\ \cos\ \phi\ \sin\phi)\ \tan^2\theta - (0{\cdot}1413\ \cos\phi + 0{\cdot}3607\ \sin\phi)\ \tan\theta - 0{\cdot}378 = 0.$$

3rd extension.

$\epsilon = 1{\cdot}381$, $f = 0{\cdot}9875$, $g = 0{\cdot}7350$, $\lambda_0 = 89{\cdot}5°$, $\lambda_1 = 87°$.
$\beta_0 = 90°\ 10'$, $\beta_1 = 82°$, $\gamma_0 = 90°$, $\gamma_1 = 112°\ 50'$.

Unstretched cone.
Unstrained position.

$$(0{\cdot}0215\ \cos^2\phi + 0{\cdot}4344\ \sin^2\phi + 0{\cdot}0078\ \cos\ \phi\ \sin\phi)\ \tan^2\theta - (0{\cdot}2965\ \cos\phi + 0{\cdot}8595\ \sin\phi)\ \tan\theta - 0{\cdot}910 = 0.$$

Strained position.

$$(0{\cdot}0409\ \cos^2\phi + 0{\cdot}9588\ \sin^2\phi - 0{\cdot}1486\ \cos\phi\sin\phi)\ \tan^2\theta - (0{\cdot}1512\ \cos\phi + 0{\cdot}4485\ \sin\phi)\ \tan\theta - 0{\cdot}478 = 0.$$

4th extension.

$\epsilon = 1{\cdot}725$, $f = 0{\cdot}9695$, $g = 0{\cdot}6510$, $\lambda_0 = 89{\cdot}5°$, $\lambda_1 = 75°$.
$\beta_0 = 90°\ 10'$, $\beta_1 = 85°\ 30'$, $\gamma_0 = 90°$, $\gamma_1 = 117°\ 30'$.

Unstretched cone.
Unstrained position.

$$(0{\cdot}0544\ \cos^2\phi + 0{\cdot}4660\ \sin^2\phi - 0{\cdot}2443\ \cos\phi\ \sin\phi)\ \tan^2\theta - (0{\cdot}2807\ \cos\ \phi + 1{\cdot}170\ \sin\phi)\ \tan\theta - 1{\cdot}975 = 0.$$

Strained position.

$$(0{\cdot}0643\ \cos^2\phi + 1{\cdot}714\ \sin^2\phi - 0{\cdot}5543\ \cos\phi\sin\phi)\ \tan^2\theta - (0{\cdot}0564\ \cos\ \phi - 0{\cdot}3798\ \sin\phi)\ \tan\theta - 0{\cdot}6645 = 0.$$

$\frac{\epsilon = 1{\cdot}725}{\epsilon = 1{\cdot}381}$ $\epsilon = 1{\cdot}247$, $f = 0{\cdot}9825$, $g = 0{\cdot}8875$, $\lambda_0 = 87°$, $\lambda_1 = 75°$.
$\beta_0 = 82°$, $\beta_1 = 85°\ 30'$, $\gamma_0 = 112°\ 50'$, $\gamma_1 = 117°\ 30'$.

Unstretched cone.
Strained position.

$$(10{\cdot}0353\ \cos^2\phi - 0{\cdot}4291\ \sin^2\phi - 0{\cdot}5511\ \cos\phi\ \sin\phi)\ \tan^2\theta + (0{\cdot}1275\ \cos\phi - 0{\cdot}1274\ \sin\phi)\ \tan\theta - 0{\cdot}358 = 0,$$

Table II—continued.

α Brass II.—Reference Face I.

1st extension.

$\epsilon = 1{\cdot}139$, $f = 0{\cdot}9500$, $g = 0{\cdot}9390$, $\lambda_0 = 88° 5'$, $\lambda_1 = 80{\cdot}5°$.
$\beta_0 = 90° 20'$, $\beta_1 = 92° 40'$, $\gamma_0 = 90° 20'$, $\gamma_1 = 82°$.

Unstretched cone.
Unstrained position.

$(0{\cdot}0956 \cos^2\phi + 0{\cdot}1051 \sin^2\phi - 0{\cdot}2403 \cos\phi \sin\phi) \tan^2\theta + (0{\cdot}0855 \cos\phi - 0{\cdot}3175 \sin\phi) \tan\theta - 0{\cdot}297 = 0.$

2nd extension.

$\epsilon = 1{\cdot}238$, $f = 0{\cdot}9395$, $g = 0{\cdot}9045$, $\lambda_0 = 88{\cdot}8°$, $\lambda_1 = 74{\cdot}2°$
$\beta_0 = 90° 20'$, $\beta_1 = 92° 50'$, $\gamma_0 = 90° 20'$, $\gamma_1 = 74° 10'$.

Unstretched cone.
Unstrained position.

$(0{\cdot}1170 \cos^2\phi + 0{\cdot}1217 \sin^2\phi - 0{\cdot}4019 \cos\phi \sin\phi) \tan^2\theta + (0{\cdot}0971 \cos\phi - 0{\cdot}6535 \sin\phi) \tan\theta - 0{\cdot}655 = 0.$

3rd extension.

$\epsilon = 1{\cdot}465$, $f = 0{\cdot}8205$, $g = 0{\cdot}9320$, $\lambda_0 = 88{\cdot}8°$, $\lambda_1 = 67{\cdot}7°$.
$\beta_0 = 90° 20'$, $\beta_1 = 99°$, $\gamma_0 = 90° 20'$, $\gamma_1 = 60° 30'$.

Unstretched cone.
Unstrained position.

$(0{\cdot}3127 \cos^2\phi - 0{\cdot}1450 \sin^2\phi - 0{\cdot}4191 \cos\phi \sin\phi) \tan^2\theta + (0{\cdot}3561 \cos\phi - 1{\cdot}5790 \sin\phi) \tan\theta - 1{\cdot}145 = 0.$

Strained position.

$(0{\cdot}6937 \cos^2\phi + 0{\cdot}8805 \sin^2\phi - 1{\cdot}3334 \cos\phi - \sin\phi) \tan^2\theta + (0{\cdot}1376 \cos\phi - 0{\cdot}6395 \sin\phi) \tan\theta - 0{\cdot}5345 = 0.$

6th extension.

$\epsilon = 1{\cdot}915$, $f = 0{\cdot}6700$, $g = 0{\cdot}9395$, $\lambda_0 = 88{\cdot}8°$, $\lambda_1 = 57°$.
$\beta_0 = 90° 20'$, $\beta_1 = 90° 30'$, $\gamma° = 90° 20'$, $\gamma_1 = 46° 30'$.

Unstretched cone.
Unstrained position.

$(0{\cdot}5510 \cos^2\phi - 0{\cdot}6795 \sin^2\phi - 0{\cdot}6740 \cos\phi \sin\phi) \tan^2\theta - (0{\cdot}0203 \cos\phi - 3{\cdot}450 \sin\phi) \tan\theta - 2{\cdot}665 = 0.$

Strained position.

$(1{\cdot}230 \cos^2\phi + 1{\cdot}8525 \sin^2\phi - 2{\cdot}7522 \cos\phi \sin\phi) \tan^2\theta - (0{\cdot}0042 \cos\phi - 0{\cdot}6215 \sin\phi) \tan\theta - 0{\cdot}7270 = 0.$

$\frac{\epsilon = 1{\cdot}915}{\epsilon = 1{\cdot}465}$ $\epsilon = 1{\cdot}308$, $f = 0{\cdot}8465$, $g = 1{\cdot}09$, $\lambda_0 = 67{\cdot}7°$, $\lambda_1 = 57°$.
$\beta_0 = 99°$, $\beta_1 = 90° 30'$, $\gamma_0 = 60° 30'$, $\gamma_1 = 46° 30'$.

Unstretched cone.
Strained position.

$(0{\cdot}4315 \cos^2\phi + 0{\cdot}3068 \sin^2\phi - 0{\cdot}7252 \cos\phi \sin\phi) \tan^2\theta - (0{\cdot}2765 \cos\phi - 0{\cdot}1221 \sin\phi) \tan\theta - 0{\cdot}415 = 0.$

$\frac{\epsilon = 2{\cdot}595}{\epsilon = 1{\cdot}915}$ $\epsilon = 1{\cdot}355$, $f = 0{\cdot}9540$, $g = 0{\cdot}9710$, $\lambda_0 = 57°$, $\lambda_1 = 42°$.
$\beta_0 = 90° 30'$, $\beta_1 = 84° 30'$, $\gamma_0 = 46° 30'$, $\gamma_1 = 30° 10'$.

Unstretched cone.
Strained position.

$(0{\cdot}1064 \cos^2\phi + 0{\cdot}8031 \sin^2\phi - 0{\cdot}6118 \cos\phi \sin\phi) \tan^2\theta - (0{\cdot}1181 \cos\phi - 0{\cdot}499 \sin\phi) \tan\theta - 0{\cdot}399 = 0.$

Table II—continued.

α Brass III.—Reference Face I.

1st extension.

$\epsilon = 1{\cdot}088$, $f = 0{\cdot}9905$, $g = 0{\cdot}9355$, $\lambda_0 = 90°$, $\lambda_1 = 84°$.
$\beta_0 = 90°\ 40'$, $\beta_1 = 90°\ 40'$, $\gamma_0 = 89°\ 10'$, $\gamma_1 = 80°\ 10'$.

Unstretched cone.
Unstrained position.

$$(0{\cdot}0188 \cos^2\phi + 0{\cdot}0172 \sin^2\phi - 0{\cdot}1940 \cos\phi \sin\phi) \tan^2\theta - (0{\cdot}0024 \cos\phi - 0{\cdot}3179 \sin\phi) \tan\theta - 0{\cdot}182 = 0.$$

2nd extension.

$\epsilon = 1{\cdot}204$, $f = 0{\cdot}9630$, $g = 0{\cdot}8780$, $\lambda_0 = 90°$, $\lambda_1 = 80{\cdot}2°$,
$\beta_0 = 90°\ 40'$, $\beta_1 = 90°\ 40'$, $\gamma_0 = 89°\ 10'$, $\gamma_1 = 70°\ 30'$.

Unstretched cone.
Unstrained position.

$$(0{\cdot}0893 \cos^2\phi - 0{\cdot}0537 \sin^2\phi - 0{\cdot}3127 \cos\phi \sin\phi) \tan^2\theta - (0{\cdot}1016 \cos\phi - 0{\cdot}7065 \sin\phi) \tan\theta - 0{\cdot}452 = 0.$$

Strained position.

$$(0{\cdot}0800 \cos^2\phi + 0{\cdot}4373 \sin^2\phi - 0{\cdot}3607 \cos\phi \sin\phi) \tan^2\theta - (0{\cdot}0398 \cos\phi - 0{\cdot}4325 \sin\phi) \tan\theta - 0{\cdot}312 = 0.$$

3rd extension.

$\epsilon = 1{\cdot}290$, $f = 0{\cdot}9345$, $g = 0{\cdot}8540$, $\lambda_0 = 90°$, $\lambda_1 = 74{\cdot}5°$.
$\beta_0 = 90°\ 40'$, $\beta_1 = 90°\ 30'$, $\gamma_0 = 89°\ 10'$, $\gamma_1 = 61°\ 30'$.

Unstretched cone.
Strained position.

$$(0{\cdot}1660 \cos^2\phi + 0{\cdot}8001 \sin^2\phi - 0{\cdot}8318 \cos\phi \sin\phi) \tan^2\theta - (0{\cdot}0089 \cos\phi - 0{\cdot}6455 \sin\phi) \tan\theta - 0{\cdot}399 = 0.$$

7th extension. Bottom.

$\epsilon = 1{\cdot}735$, $f = 0{\cdot}8595$, $g = 0{\cdot}7930$, $\lambda_0 = 90°$, $\lambda_1 = 57{\cdot}5°$.
$\beta_0 = 90°\ 40'$, $\beta_1 = 66°\ 30'$, $\gamma_0 = 89°\ 10'$, $\gamma_1 = 48°\ 40'$.

Unstretched cone.
Strained position.

$$(0{\cdot}4184 \cos^2\phi + 0{\cdot}8411 \sin^2\phi - 1{\cdot}5015 \cos\phi \sin\phi) \tan^2\theta - (0{\cdot}3108 \cos\phi - 0{\cdot}4705 \sin\phi) \tan\theta - 0{\cdot}667 = 0.$$

7th Extension. Top.

$\epsilon = 2{\cdot}685$, $f = 0{\cdot}8766$, $g = 0{\cdot}7095$, $\lambda_0 = 90°$, $\lambda_1 = 37°$.
$\beta_0 = 90°\ 40'$, $\beta_1 = 50°$, $\gamma_0 = 89°\ 10'$, $\gamma_1 = 25°\ 10'$.

Unstretched cone.
Strained position.

$$(0{\cdot}4095 \cos^2\phi + 7{\cdot}579 \sin^2\phi - 2{\cdot}901 \cos\phi \sin\phi) \tan^2\theta - (0{\cdot}2425 \cos\phi - 0{\cdot}6340 \sin\phi) \tan\theta - 0{\cdot}8615 = 0.$$

Extension, 7th to 5th.

$$\frac{\epsilon = 1{\cdot}735}{\epsilon = 1{\cdot}581} \quad \epsilon = 1{\cdot}096, \quad f = 0{\cdot}9620, \quad g = 0{\cdot}9885, \quad \lambda_0 = 63°, \quad \lambda_1 = 57{\cdot}5°.$$
$$\beta_0 = 77°, \quad \beta_1 = 66°\ 30', \quad \gamma_0 = 54°, \quad \gamma_1 = 48°\ 40'.$$

Unstretched cone.
Strained position.

$$(0{\cdot}1074 \cos^2\phi + 0{\cdot}1601 \sin^2\phi - 0{\cdot}2537 \cos\phi \sin\phi) \tan^2\theta - (0{\cdot}3025 \cos\phi + 0{\cdot}0345 \sin\phi) \tan\theta - 0{\cdot}167 = 0.$$

Table II—(continued).

α Brass III.—Reference Face I—continued.

7th extension. Top to bottom.

$\epsilon = 2{\cdot}685$
$\epsilon = 1{\cdot}735$ $\epsilon = 1{\cdot}545$, $f = 1{\cdot}019$, $g = 0{\cdot}8940$, $\lambda_0 = 57{\cdot}5°$, $\lambda_1 = 37°$.

$\beta_0 = 66° 30'$, $\beta_1 = 50°$, $\gamma_0 = 48° 40'$, $\gamma_1 = 25° 10'$.

Unstretched cone.

Strained position.

$$(0{\cdot}0206 \cos^2\phi - 1{\cdot}7937 \sin^2\phi + 0{\cdot}4784 \cos\phi \sin\phi) \tan^2\theta + (0{\cdot}1510 \cos\phi + 0{\cdot}6490 \sin\phi) \tan\theta + 0{\cdot}581 = 0.$$

Particulars as to methods of calculation, etc., are to be found in the paper just referred to, and as the notation is the same throughout this paper and that on the aluminium-zinc alloys no further explanation is necessary. Only the values necessary for the calculation of the cone of unextended directions, and the equations of the cone are given. At first all calculations were made in reference to $\epsilon = 1{\cdot}0$. Later, however, calculations were made between different stages and sections of the extension, as from the behaviour of the test-pieces it became evident that the distortion was of a complicated nature. The most important of these are given in Table II where it is shown between which stages the calculations were made.

A complete set of measurements were made on three separate brass crystals, but other specimens have been tested which show the same characteristics. As the history of each test-piece is different they will be described separately.

Crystal I.

The position of the axis of the test-piece in relation to the crystal is shown in fig. 1.* The first small extension (about 2 per cent.) produced a number of slip-bands, figs. 1 and 2, Plate 3, of which one set was clearly the most important, and this set persisted and increased in number as the test proceeded. The appearance of more than one set of slip-lines at the beginning of a test has been observed by Gough, Hanson and Wright in aluminium test-pieces subjected to alternating stress tests.† It is a little uncertain how much this is due to the remains of strain in the surface caused by machining and polishing, as there is no means of ascertaining if these planes go right through the test-piece. They may account for small discrepancies between results in distortion calculations and X-ray measurements which will be referred to later. The

* The means whereby this diagram is obtained and made use of has been described in previous papers. In order to avoid confusion, the same form of projection is used here. G. I. Taylor and C. F. Elam, ' Roy. Soc. Proc.,' A, vol. 108 (1925).

† Gough, Hanson and Wright, ' Phil. Trans.,' A, vol. 206.

main set of slip-bands were not regularly spaced as fig. 1, Plate 3, shows, but as the extension proceeded, and they became more numerous, the gaps

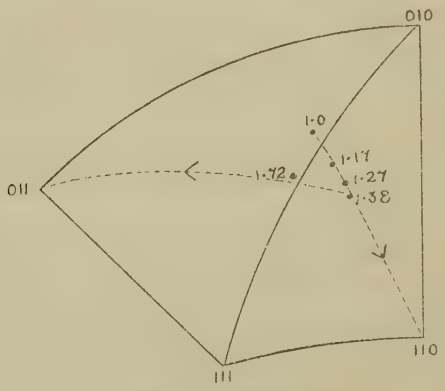


Fig. 1.—α Brass I.

closed up. It will be seen from fig. 1 that this crystal might have been expected to slip on one octahedral plane for a short time, up to about 4 per cent. extension, and after that, simultaneously on two octahedral planes, provided it behaved as copper and aluminium. X-ray measurements made at about 17 per cent. extension, however, showed that the test-piece axis had crossed the line made by the great circle joining the poles of the (010) and (111) planes where it would be equally inclined to both the (11$\bar{1}$) and ($\bar{1}$11) planes, and at which point double slipping should have begun, and reached the point marked on the diagram, well inside the adjacent triangle. Distortion measurements showed that the deformation had occurred by slip on one plane, the ($\bar{1}$11) plane according to the projection made use of in fig. 1, and towards the pole of the (110) plane. Both the slip-plane and direction of slip are those predicted from previous experience, and in the work on aluminium* it had been noticed that slip tended to go a little further on the first octahedral plane than was required by theory, but this was small compared with the present instance.

In Table III are given the positions of the slip-plane and direction of slip obtained from distortion and X-ray measurements. The agreement is not so good as that obtaining between similar measurements for aluminium and aluminium-zinc alloys, but is sufficiently close to warrant the conclusion that slip is on the particular plane and in the direction stated. The specimens were rather small for accurate measurement, and it has already been pointed out that errors in distortion may have been introduced by small amounts of slip

* G. I. Taylor and C. F. Elam, 'Roy. Soc. Proc.,' A, vol. 108 (1925). See also Gough, Hanson and Wright, *loc. cit.*

on other planes at the beginning of the test. Moreover, the chief differences are in the values of ϕ, which are most difficult to measure accurately.

Table III.

α Brass I.—Reference Face I.

Slip-plane.		Direction of slip.	
X-ray Measurements.	Distortion.	X-ray measurements.	Distortion.
$\epsilon = 1{\cdot}175$. Unstrained			
$\theta_{\bar{1}11} = 55° 40'$	$\theta = 53° 10'$	$\theta_{110} = 37° 10'$	$\theta = 37° 0'$
$\phi = 279° 30'$	$\phi = 285° 0'$	$\phi = 85° 30'$	$\phi = 91° 30'$
Strained			
$\theta_{\bar{1}11} = 58° 0'$	$\theta = 60° 0'$	$\theta_{110} = 32° 0'$	$\theta = 32° 0'$
$\phi = 276° 30'$	$\phi = 275° 0'$	$\phi = 82° 40'$	$\phi = 75° 30'$
$\epsilon = 1{\cdot}272$. Strained			
$\theta_{\bar{1}11} = 62° 0'$	$\theta = 62° 0'$	$\theta_{110} = 29° 0'$	$\theta = 28° 40'$
$\phi = 276° 0'$	$\phi = 278° 20'$	$\phi = 81° 30'$	$\phi = 81° 30'$
$\epsilon = 1{\cdot}381$. Strained			
$\theta_{\bar{1}11} = 64° 10'$	$\theta = 63° 30'$	$\theta_{110} = 27° 20'$	$\theta = 26° 0'$
$\phi = 285° 30'$	$\phi = 278° 30'$	$\phi = 92° 30'$	$\phi = 89° 20'$
$\epsilon = 1{\cdot}725$. Strained			
$\theta_{\bar{1}11} = 57° 10'$	$\theta = 62° 40'$	$\theta_{110} = 30° 30'$	$\theta = 28° 0'$
$\phi = 271° 30'$	$\phi = 280° 0'$	$\phi = 89° 0'$	$\phi = 95° 30'$
$\epsilon = 1{\cdot}725$ $\theta_{11\bar{1}} = 58° 0'$	$\theta = 56° 0'$	$\theta_{011} = 32° 0'$	$\theta = 32° 20'$
$\epsilon = 1{\cdot}381$ $\phi = 123° 0'$	$\phi = 119° 0'$	$\phi = 306° 0'$	$\phi = 288° 0'$

α Brass II.—Reference Face I.

Slip-plane.		Direction of slip.	
X-ray measurements.	Distortion.	X-ray measurements.	Distortion.
Strained			
$\epsilon = 1{\cdot}465$ $\theta_{\bar{1}11} = 68° 50'$	$\theta = 68° 0'$	$\theta_{110} = 28° 0'$	$\theta = 29° 0'$
$\epsilon = 1{\cdot}0$ $\phi = 118° 40'$	$\phi = 120° 30'$	$\phi = 339° 30'$	$\phi = 341° 30'$
$\epsilon = 1{\cdot}915$ $\theta_{\bar{1}1\bar{1}} = 64° 0'$	$\theta = 58° 0'$	$\theta_{011} = 39° 46'$	$\theta = 40° 0'$
$\epsilon = 1{\cdot}465$ $\phi = 347° 30'$	$\phi = 337° 30'$	$\phi = 110° 0'$	$\phi = 116° 30'$
$\epsilon = 2{\cdot}595$ $\theta_{\bar{1}11} = 69° 40'$	$\theta = 67° 0'$	$\theta_{110} = 26° 0'$	$\theta = 30° 0'$
$\epsilon = 1{\cdot}915$ $\phi = 102° 30'$	$\phi = 98° 40'$	$\phi = 326° 0'$	$\phi = 321° 0'$

Table III—(continued).

α Brass III.—Reference Face I.

Slip-plane.		Direction of slip.	
X-ray measurements.	Distortion.	X-ray measurements.	Distortion.
Strained			
$\epsilon = 1{\cdot}204$ $\theta_{\bar{1}11} = 68° 30'$	$\theta = 65° 0'$	$\theta_{110} = 28° 30'$	$\theta = 30° 0°$
$\epsilon = 1{\cdot}0$ $\phi = 99° 30'$	$\phi = 100° 0'$	$\phi = 320° 0'$	$\phi = 317° 30'$
$\epsilon = 2{\cdot}685$ $\theta_{\bar{1}11} = 69° 28'$	$\theta = 68° 40'$	$\theta_{110} = 24° 10'$	$\theta = 25° 30'$
$\epsilon = 1{\cdot}735$ $\phi = 85° 30'$	$\phi = 91° 0'$	$\phi = 302° 10'$	$\phi = 300° 0'$

Measurements when $\varepsilon = 1{\cdot}272$ showed that slip had occurred still on the same plane and in the same direction. (See Table III and fig. 1.) Only the principal set of slip-bands of the first extension increased in number and these agreed with the slip-plane as determined from distortion and X-ray measurements. Fig. 2 is a stereographic diagram showing the slip-plane the pole of

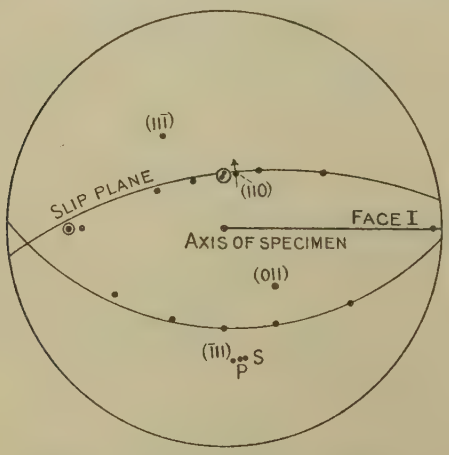


Fig. 2.—α Brass I. $\epsilon = 1{\cdot}272$. Strained position.

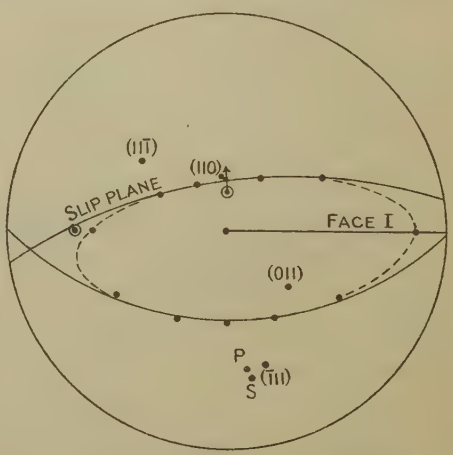


Fig. 3.—α Brass I. $\epsilon = 1{\cdot}381$. Strained position.

which is marked P, the second undistorted plane, the direction of slip marked by an arrow —→, and the crystal planes nearest to these points. The traces of slip-bands are marked also, ⊙, and the pole of this plane S. The same marks are used throughout the diagrams. Diagrams such as this were made for all stages of the deformation but only a few representative ones are given.

The indices chosen to represent the planes are the same in all the diagrams and have no special relation to the crystal structure.

At an extension of $\varepsilon = 1{\cdot}381$, slip still appeared to be on the same plane, but traces of a second slip-plane were visible. Up to this point, the points representing the test-piece axis lie on the great circle joining the point of origin to the pole of the (110) plane in fig. 1, adding confirmatory evidence that slip had occurred on the $(\bar{1}11)$ plane and in the (110) direction. If slip had at any time begun on the $(11\bar{1})$ plane, the direction of slip on this plane would have been towards the pole of the (011) plane. In the case of aluminium it was shown that by slip first on the $(\bar{1}11)$ plane in the (110) direction, and then on the $(11\bar{1})$ plane in the (011) direction, the axis of the specimen gradually moved along the great circle towards the pole of the (121) plane, when no further change would be expected, as this point lies on the great circle joining both directions of slip. The factor chiefly responsible for this type of deformation is the shear-stress on the slip-plane. As soon as the axis has crossed the boundary, the shear-stress on the second octahedral plane is greater than on the first, and slip then occurs on this plane, and the axis tends to move back across the boundary. This process is repeated with first one plane and then the other, with the result already described.

The brass crystal having reached the position shown in fig. 1, the shear-stress on the second octahedral plane, the $(11\bar{1})$ plane, was much greater than on the first,* and at some point near the extension $\varepsilon = 1{\cdot}381$, slip must have begun on this plane, and in a direction represented by the pole of the (011) plane. The change over from one slip-plane to another is easily seen on the specimen. It sometimes starts in one place, or sometimes at two or three points simultaneously, and it is some time before the specimen appears to be uniform once more. In the case of Crystal I the next measurements were made at an extension of 72 per cent. ($\varepsilon = 1{\cdot}725$). X-ray analysis showed that the axis had moved towards the (011) plane, and was almost at the point for double slipping, but was back in the first triangle as fig. 1 shows. The second set of slip-lines, which were just visible at $\varepsilon = 1{\cdot}381$, were the dominant set, the first having nearly disappeared. The specimen had become rather tarnished at this stage, and after cleaning with a little dilute ammonia, it was found that, while the second set of slip-lines were unaffected, the first set were entirely removed. Calculation of the cone of unextended directions from measurements corresponding to $\varepsilon = 1{\cdot}381$ and $\varepsilon = 1{\cdot}725$, showed that the points did not lie very accurately on two planes, but taking the nearest great circle through them,

* See later, page 163.

the pole of one plane was very close to the pole of the $(11\bar{1})$ plane. (Table III.) The direction of slip, which could only be obtained approximately, did not agree very closely with the pole of the (011) plane, but fig. 1 shows that the main direction of slip must have been towards this point. It is also probable that slip may have occurred on both these planes at some time during the period, and this view is supported by the fact that when the specimen was stretched further, it broke almost immediately, showing a double-wedge fracture similar to those of copper and aluminium, which it has been shown, is due to slipping on two planes. The final stage of this crystal, therefore, resembles the pure metals of the same crystal structure, and unless intermediate measurements had been made, it would have appeared quite normal.

Crystal II.

The position of this crystal relative to the axis of the specimen is shown in fig. 4. It was calculated that this crystal would slip on the $(\bar{1}11)$ plane until ε was equal to 1·29. Calculations made from measurements when $\varepsilon = 1{\cdot}139$ and $\varepsilon = 1{\cdot}238$, confirmed this view. In order not to confuse the diagram the position of the axis at these points is not included in fig. 4. Calculations when $\varepsilon = 1{\cdot}465$ still showed slip on the same plane and in the same direction. This point is shown in fig. 4, and in Table III are given the positions of the slip-plane and direction of slip obtained from X-ray and distortion measurements. Up to this point only one set of slip-planes was visible, fig. 3, Plate 3, except for traces of other sets which appeared previously. Fig. 5 is a stereographic diagram representing the distortion at this point.

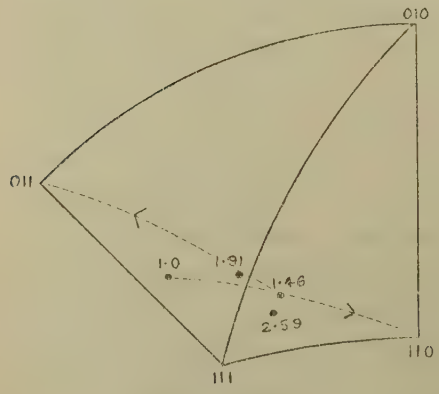


Fig. 4.—α Brass II.

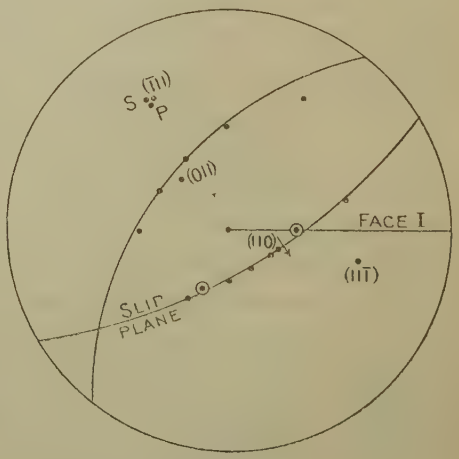


Fig. 5.—α Brass II. $\epsilon = 1{\cdot}46$. Strained position.

On replacing the specimen in the testing machine, and continuing the extension, it suddenly appeared to give way, first at one end, and then at a few points elsewhere. It was removed from the machine and photographed (fig. 4, Plate 3, and fig. 1, Plate 4). A second set of very well-marked slip-bands appeared (fig. 4, Plate 3). The specimen was too uneven to give uniform measurements, so it was stretched further. As soon as the distortion appeared to be even again, the crystal started narrowing abruptly near one end. (The specimen was slightly tapered, so that changes began near the small end.) Instead of breaking, as it appeared at first, a third type of deformation began. When nearly half of the specimen had changed it was carefully measured. A photograph is shown in fig. 3, Plate 4.

X-ray measurements were made on both parts of the specimen. The bottom part of the crystal had an extension represented by the value $\varepsilon = 1{\cdot}91$, the top that of $\varepsilon = 2{\cdot}595$. These two points are shown in fig. 4. During the period from $\varepsilon = 1{\cdot}46$ to $\varepsilon = 1{\cdot}91$, the axis moved towards the pole of the (011) plane, while from $\varepsilon = 1{\cdot}91$ to $\varepsilon = 2{\cdot}595$ it tended to move again towards the (110) plane. Distortion calculations between these extensions, viz., from $\varepsilon = 1{\cdot}46$ to $\varepsilon = 1{\cdot}91$, showed that the points did not lie accurately on two planes, fig. 6, but drawing the nearest great circle through them, the pole was near the $(11\bar{1})$ plane, and the direction of slip near the pole of the (011) plane (Table III). There appeared to be no connection with the $(\bar{1}11)$ plane. This also agreed with the results of slip-band measurement (fig. 6). Of the two sets of slip-bands now on the surface the new set agreed more closely with the $(11\bar{1})$ plane than the first set did with the $(\bar{1}11)$ plane. If calculations were made between measurements at $\varepsilon = 1{\cdot}0$ and $\varepsilon = 1{\cdot}91$ a cone-shaped figure was obtained indicating slip on two planes.

Calculations were then made between measurements obtained from the top and bottom of the specimen, *i.e.*, between values of $\varepsilon = 1{\cdot}91$ and $\varepsilon = 2{\cdot}595$. The points fell more closely on the great circles, the pole of one plane now agreeing with the pole of the $(\bar{1}11)$ plane, *i.e.*, with the first slip-plane, although the agreement between the direction of slip and the (110) plane was not so good (fig. 7). The figures are given in Table III. Traces of the third set of slip-lines also agreed with this plane. Traces of the first two sets of slip-lines remained on the surface, but now had no more connection with the crystal structure than any other scratches or marks on the surface. This only tends to emphasise the caution that is necessary in basing any conclusions as to the plane of slip, on the measurements of slip-bands alone. They frequently do agree with the plane of slip, but only while it is actually functioning as the

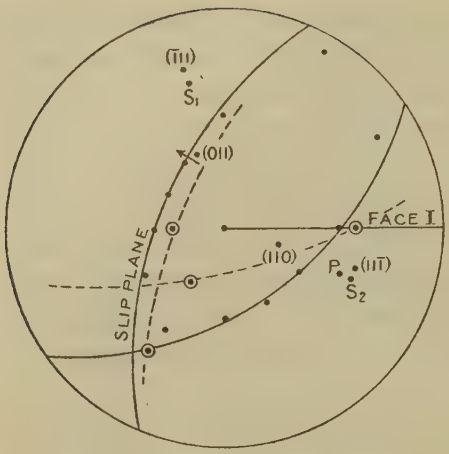


Fig. 6.—α Brass II. $\epsilon = 1{\cdot}465$ to $\epsilon = 1{\cdot}915$. Strained position.

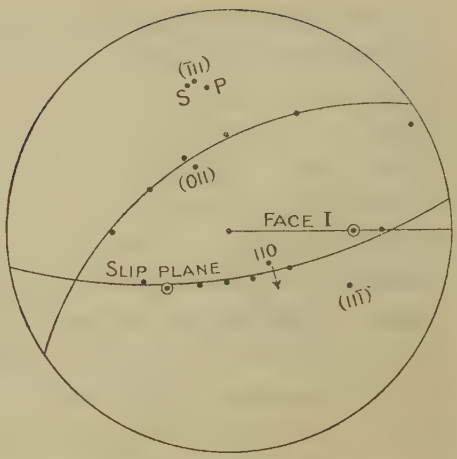


Fig. 7.—α Brass II. $\epsilon = 1{\cdot}915$ to $\epsilon = 2{\cdot}595$. Strained position.

slip-plane. It is possible, however, to separate the true from the false, by repolishing or re-etching the specimen at intervals.

Fig. 8 is a drawing to scale representing the original and final cross-sections of this crystal. It gives a very good idea of the enormous distortion this crystal

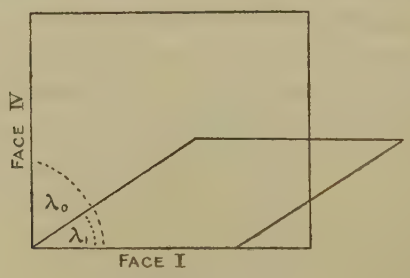


Fig. 8.

has undergone, and it also shows the difficulty of obtaining accurate measurements of the width of the faces and the angle λ.

The distortion of this crystal, therefore, can really be divided into three stages, as its behaviour during testing indicated. Slip was on the same plane, and approximately in the same direction in the first and third stages, but on a different plane, and in a reverse direction during the second stage. Like the first crystal, slip occurred on the first plane over a much longer extension than in the pure metals. On extending the crystal still further, the third stage spread down the test-piece, but the specimen then broke with a straight fracture perpendicular to the axis, in the narrowest part, without any further extension.

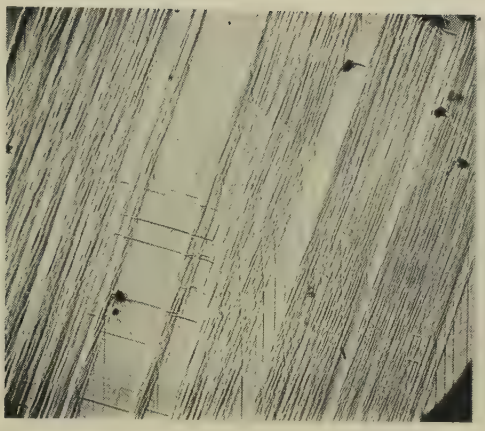


FIG. 1.—Crystal I. 1st Extension, Face I.

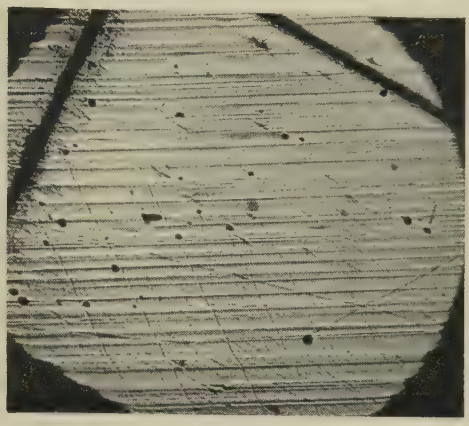


FIG. 2.—Crystal I. 1st Extension, Face II.

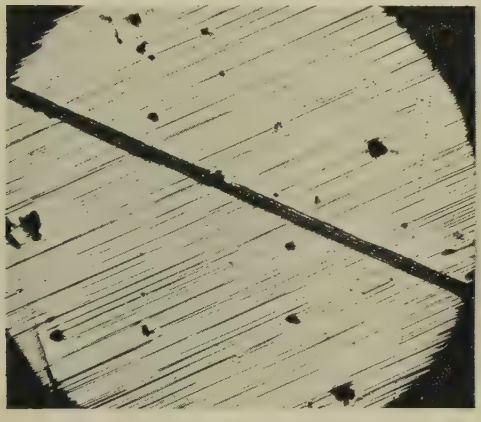


FIG. 3.—Crystal II. 1st Extension.

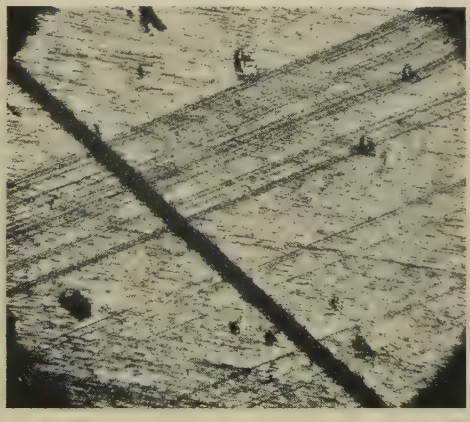


FIG. 4.—Crystal II. 4th Extension.

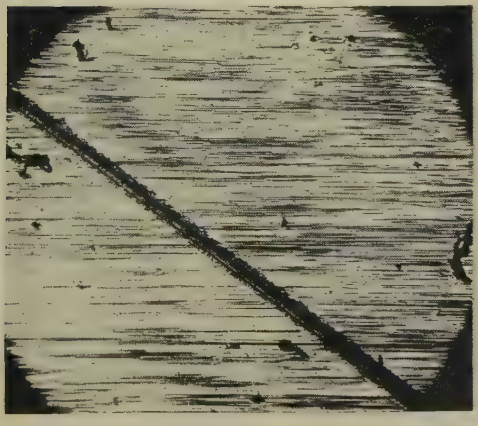


FIG. 5.—Crystal III. 1st Extension.

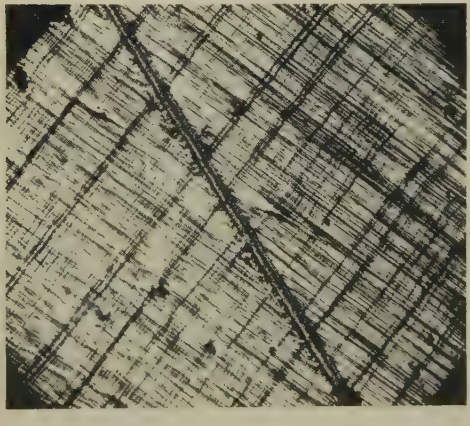


FIG. 6.—Crystal III. 4th Extension.

Magnification—45 diameters.

(*Facing p.* 160.)

FIG. 1.
Crystal II at beginning of 2nd stage.

FIG. 2.
Crystal III at beginning of 2nd stage.

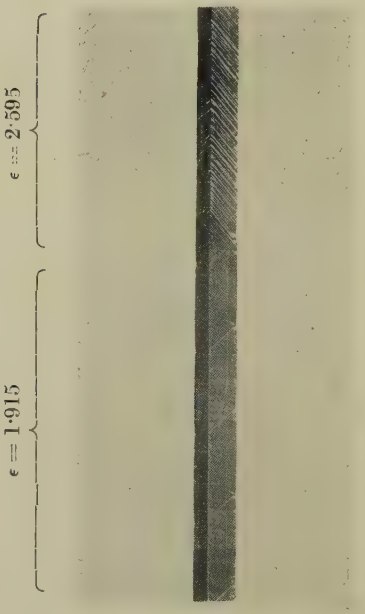


FIG. 3.
Crystal II, last stage.

FIG. 4.
Crystal III, last stage.

Crystal III.

The original position of this crystal is shown in fig. 9. Slip on two planes should have begun at about 5 per cent. extension, but distortion and X-ray measurements showed that at extensions corresponding to $\varepsilon = 1{\cdot}088$, and $\varepsilon = 1{\cdot}204$, slip was still on the first octahedral plane, and towards the (110) plane. Only one set of slip-bands was visible (figs. 9 and 10, and Table III,

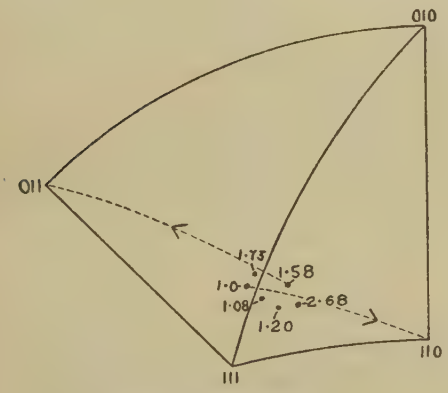


Fig. 9.—α Brass III.

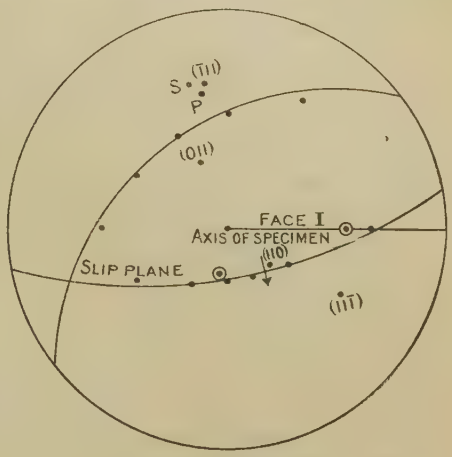


Fig. 10.—α Brass III. $\epsilon = 1{\cdot}204$. Strained position.

and fig. 5, Plate 3.) On extending further, however ($\varepsilon = 1{\cdot}292$), two sets of slip-bands were visible, and calculations showed that the cone of unextended directions no longer consisted of two planes. The position of the test-piece axis was almost unaltered. At this point the test-piece became very uneven, a new type of deformation setting in at several points, fig. 2, Plate 3. At a further extension ($\varepsilon = 1{\cdot}581$) the test-piece became more uniform and measurements were made (fig. 6, Plate 3). The position of the axis is shown in fig. 9. It had moved parallel to the great circle through the poles of the (010) and (111) planes, but at some distance from it, indicating that slip had occurred on both ($\bar{1}11$) and ($11\bar{1}$) planes, but the axis was not symmetrical to these planes. The test-piece was still not quite uniform, so it was extended further. As in crystal No. II, no sooner did it appear uniform than it suddenly drew down near one end, thus starting a third stage in the deformation. When about one-third of the test-piece had changed, it was removed from the machine and both parts carefully measured. In the least extended part ε was equal to 1·735, and the position of the test-piece axis is shown in fig. 9. The point had re-crossed the boundary for double slipping and was nearer to the pole of the

(121) plane. Judging from this, most of the distortion probably occurred on the (11$\bar{1}$) plane, towards the pole of the (011) plane, but this could not be proved as calculations of the distortion, from $\varepsilon = 1{\cdot}581$ to $\varepsilon = 1{\cdot}735$, gave a cone indicating that slip had been on both planes (fig. 11). Traces of slip-bands, however, agreed with this plane. Calculations between the two last

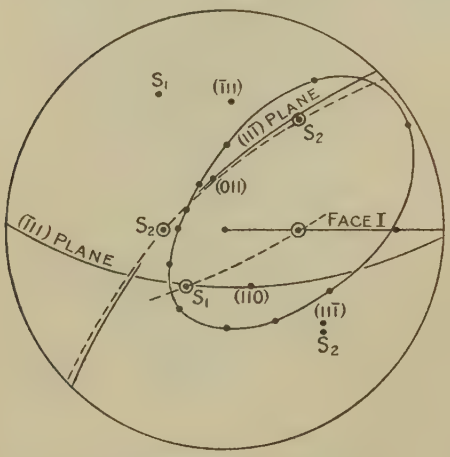


Fig. 11.—α Brass III. $\epsilon = 1{\cdot}581$ to $\epsilon = 1{\cdot}735$. Strained position.

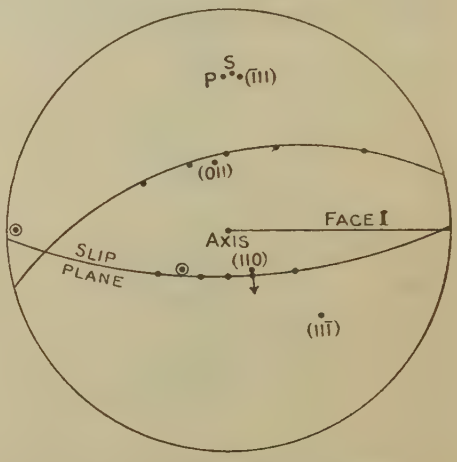


Fig. 12.—α Brass III. $\epsilon = 1{\cdot}735$ to $\epsilon = 2{\cdot}685$. Strained position.

stages of the distortion agreed closely with slip on the original (1̄11) plane in the direction of the (110) plane (Table III, fig. 12). Traces of slip-bands also agreed with this plane, and X-ray analysis showed that the axis had moved in a direction conforming with this conclusion (fig. 9).

As in crystal II the deformation took place in three stages of which the first and last were on the same plane and in the same direction, while the middle stage appears to have been made up of irregular slipping on both these planes. The crystal broke in the thinnest part before the final stage had been completed. Fig. 4, Plate 4, shows its appearance just broken, and brings out the abrupt change in cross-section as the deformation passes from one form to the other.

The most noticeable features of the distortion of these crystals are, firstly, the enormous extension, *i.e.*, 160 per cent., before fracture, compared with copper which rarely exceeds 60 per cent., and, secondly, the uniformity of the crystal in its final condition, considering the abrupt changes in shape when slip changes from one plane to the other. At this point it appears as if the planes must be considerably bent, yet they straighten out again without any irregularities. This type of deformation very much resembles that described

by Mark, Polanyi and Schmid* as "Nachdehnung," as applied to a similar appearance in the deformation of zinc. These authors definitely co-relate this structure with a bending of the crystal planes. A similar effect is also to be found at the ends of single crystal test-pieces, at the junction between distorted and undistorted metal.

Calculations of the shear-stresses on both planes were made. The results are given in Table IV and plotted in figs. 13 and 14. Some curves obtained from copper crystals are also shown. θ is the angle between the axis of the specimen and the normal to the slip-plane, and η the angle between the direction of slip and the direction of greatest slope. These values are obtained from X-ray measurements using a stereographic net. S_1 and S_2 are the shear-stresses in pounds per square inch, in directions parallel to the first and second slip-planes respectively. It is to be observed that where the slip is definitely on one plane or the other, the points lie on a straight line and that the shear-stress continues to increase on a plane when it ceases to be the slip-plane at the same rate as it did when it functioned as the slip-plane. When slip begins on a plane, which has not been acting as the slip-plane, the shear-stress decreases on that plane. The shear-stress on both planes is equal at certain points, corresponding to positions where simultaneous double slipping would normally take place. It will also be noticed that the rate of increase in shear-stress is much less the second time that slip occurs on the plane. This is very well brought out in the case of Crystal II, and, to a less extent, in Crystal III. There appears to be no relation between the shear stresses on the two planes at the point when one ceases to act as slip-plane and the other begins. Where slip occurs on both planes simultaneously, the rate of increase of resistance to shear is more rapid than when either plane acts separately. This is brought out in the case of Crystal III shown in fig. 14, where it will be seen that the slope of the curve in the middle section, *i.e.*, from $\varepsilon = 1{\cdot}204$ to $\varepsilon = 1{\cdot}735$, is steeper than in either the first or last stages. Copper crystals show the same effect.

The properties of these alloy crystals, just described, suggests that the different octahedral planes in the crystal are not all equally resistant to shear. The most probable explanation is that owing to their method of manufacture the crystals were not homogeneous. As the crystal solidified it is probable that alternate layers of zinc-rich and copper-rich material were deposited on certain crystal planes, and that even after a long period of heating, diffusion

* Mark, Polanyi, Schmid, 'Z. f. Physik' (1922).

Table IV.

ϵ.	First slip-plane. ($\bar{1}11$).		Second slip-plane ($11\bar{1}$)		Cross-sectional area.	Load.	Shear stress.	
	θ.	η.	θ.	η.			S_1.	S_2.
Crystal I—					sq. ins.	lbs.	lbs. per	sq. in.
1·00	55° 40′	8° 30′	58° 30′	11° 40′	—	—	—	—
1·17	59° 0′	7° 0′	55° 40′	5° 0′	0·0252	230	4,001	4,225
1·27	62° 0′	6° 0′	53° 0′	4° 50′	0·0233	380	6,710	7,710
1·381	64° 0′	6° 0′	50° 30′	1° 50′	0·0213	500	9,025	11,660
1·725	57° 10′	0° 45′	58° 0′	1° 30′	0·0169	520	13,980	13,780
Crystal II—								
1·0	55° 30′	26° 0′	75° 0′	15° 10′	—	—	—	—
1·235	60° 30′	19° 30′	66° 40′	17° 0′	0·0160	187	4,700	4,055
1·465	68° 50′	14° 10′	59° 30′	14° 40′	0·0131	381	9,450	12,240
1·915	61° 0′	16° 20′	64° 30′	18° 0′	0·0100	455	18,430	16,780
2·595	69° 40′	16° 30′	60° 50′	23° 30′	0·0075	455	18,860	23,550
Crystal III								
1·0	62° 30′	18° 0′	69° 30′	17° 30′	—	—	—	—
1·088	68° 30′	22° 50′	64° 50′	20° 30′	0·0196	150	2,410	2,760
1·204	68° 30′	16° 50′	61° 30′	22° 20′	0·0177	235	4,235	5,145
1·292	68° 30′	16° 50′	61° 30′	22° 20′	0·0162	369	7,415	8,710
1·581	62° 30′	8° 0′	60° 0′	16° 40′	0·0126	430	13,880	14,220
1·695	60° 48′	15° 10′	65° 0′	13° 20′	0·0122	470	15,760	14,290
1·735	61° 20′	14° 30′	64° 40′	14° 0′	0·0119	490	16,710	15,360
2·685	69° 28′	12° 20′	58° 30′	21° 0′	0·0079	490	19,880	25,750

	ϵ.	Load.	Area.	ϑ.	η.	S.
Copper XI		lbs.	sq. ins.			lbs. sq. ins.
	1·00	—	0·0492	42° 03′	18° 0′	—
	1·10	287	0·0447	47° 0′	12° 40′	3,230
	1·25	683	0·0373	55° 30′	10° 0′	8,430
	1·30	764	0·0341	55° 46′	10° 0′	10,250
	1·40	815	0·0317	58° 40′	10° 40′	11,200
	1·52	857	0·0278	63° 20′	8° 40′	12,220
Copper X						
	1·00	—	0·0490	62° 0′	0° 0′	—
	1·10	512	0·0450	—	—	4,665
	1·20	810	0·0399	—	—	8,325
	1·30	934	0·0370	—	—	10,380
	1·40	979	0·0327	—	—	12,270
	1·50	987	0·0298	—	—	13,580

Crystal I.—Slip is on 1st Slip-plane ($\bar{1}11$) from $\epsilon = 1{\cdot}0$ to $\epsilon = 1{\cdot}464$.
Slip is on 2nd Slip-plane ($11\bar{1}$) from $\epsilon = 1{\cdot}915$ to $\epsilon = 1{\cdot}725$.
Crystal II.—Slip is 1st Slip-plane from $\epsilon = 1{\cdot}0$ to $\epsilon = 1{\cdot}465$.
Slip is 1st Slip-plane and from $\epsilon = 1{\cdot}915$ to $\epsilon = 2{\cdot}595$.
Slip is on 2nd Slip-plane from $\epsilon = 1{\cdot}465$ to $\epsilon = 1{\cdot}915$.
Crystal III.—Slip is on 1st Slip-plane from $\epsilon = 1{\cdot}0$ to $\epsilon = 1{\cdot}204$ and from $\epsilon = 1{\cdot}735$ to $\epsilon = 2{\cdot}685$.
Between $\epsilon = 1{\cdot}204$ and $\epsilon = 1{\cdot}735$ slip occurred on both planes, but from $\epsilon = 1{\cdot}581$ to $\epsilon = 1{\cdot}735$ slip occurred mostly on second slip-plane.

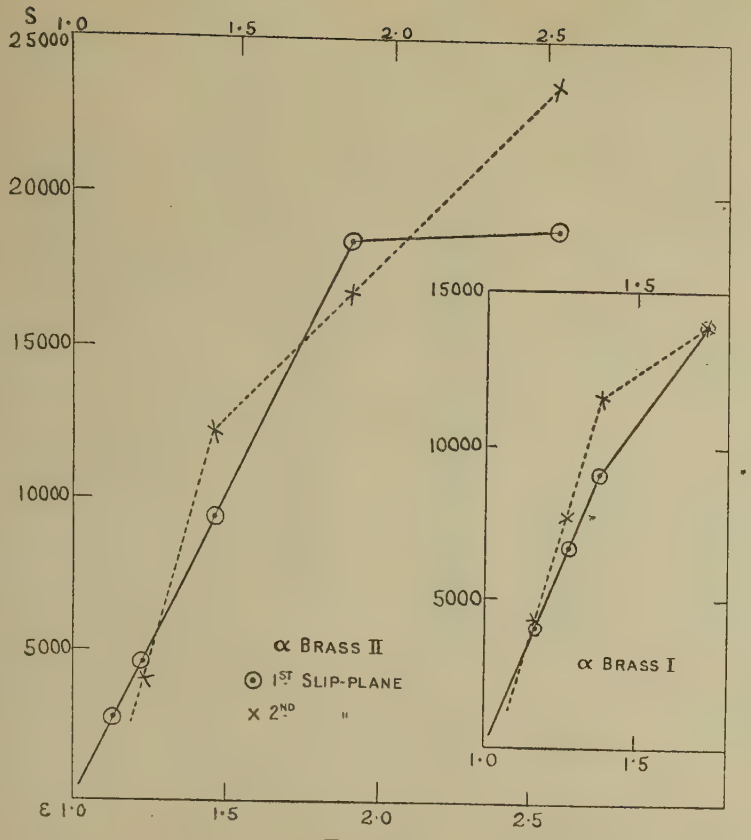

S
25000
20000
15000
10000
5000
1·0
1·5
2·0
2·5
ε
α Brass II
⊙ 1st Slip-plane
× 2nd "
α Brass I

Fig. 13.

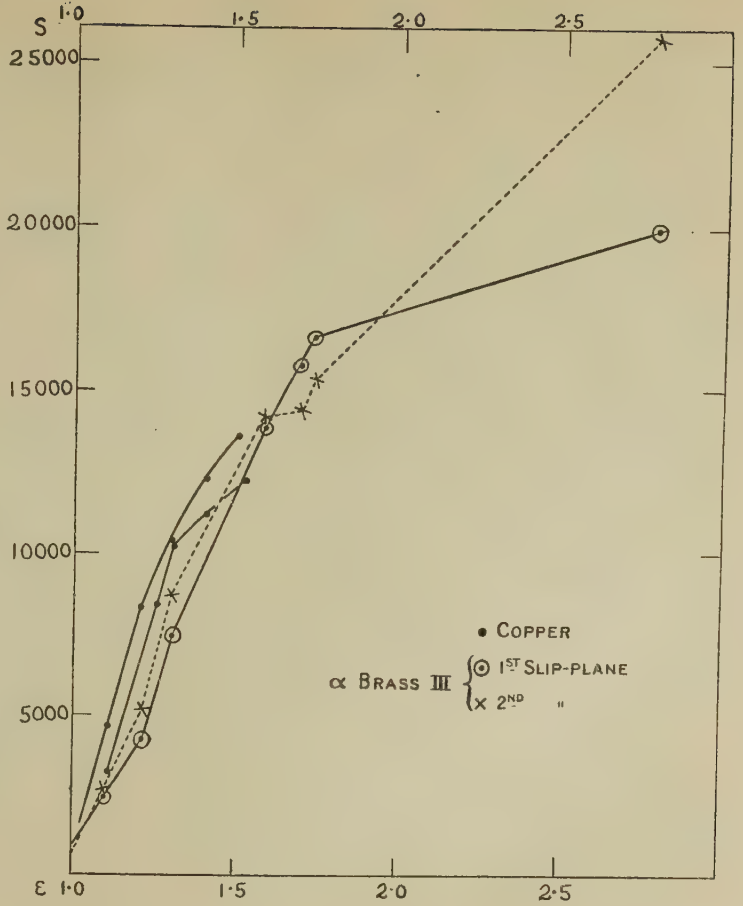

S
25000
20000
15000
10000
5000
1·0
1·5
2·0
2·5
ε
• Copper
α Brass III
⊙ 1st Slip-plane
× 2nd "

Fig. 14.

was incomplete. It is well known that diffusion without mechanical work is extremely slow. Evidence in favour of this hypothesis is that the slip-lines tend to be in bands, with intermediate regions almost free at the beginning of deformation. On the other hand, distortion on both octahedral planes shows wide bands. Moreover, the orientations of the three crystals investigated differ widely, and it appears improbable that the second slip-plane should in all three cases contain more zinc atoms, just by chance. Again, in both Crystals I and III, there are two octahedral planes almost equally inclined to the axis, so that it might be expected that both would be similar as regards solidification, yet slip still occurs, first on that plane which, according to all that is known of the distortion of metal crystals, should be the slip-plane. Therefore, although this is a possible explanation of the different behaviour of otherwise similar crystallographic planes, it does not seem to be the most probable in this particular case.

It appears rather that slip tends to proceed more easily along a plane upon which it has once begun although this is contrary to the usually accepted view that gliding on certain planes is more difficult in the case of an alloy than in that of the pure metal, owing to the distortion of the crystal lattice due to the presence of the alloying element.* In fig. 14 are plotted two curves obtained from similar measurements on copper crystals. It is interesting to find that there is so little difference between these and the brass crystals, although the latter are finally much harder. The difference in the amount of elongation before fracture is also very marked.† This seems in some way to be connected with the fact that a much greater extension is obtained, the longer slip is confined to one plane. These results, however, are the opposite to those obtained with the aluminium-zinc alloys.

* Rosenhain, ' Jour. Inst. Metals,' vol. 30 (1923), May Lecture.

† Normal annealed α brass has usually both a higher tensile strength and elongation than pure copper.

Tensile Tests on Alloy Crystals. Part III—*Conclusions.*

By C. F. Elam, Armourers' and Braziers' Research Fellow.

(Communicated by H. C. H. Carpenter, F.R.S.—Received March 29, 1927.)

It is now generally accepted that with one or two exceptions when one metal is added to another and the two form a "solid solution" or mixed crystals, it is brought about by the substitution of certain atoms of the principal metal by those of the alloying element. The crystal-lattice expands or contracts according to the nature of the alloying metal. It has not yet been shown that this substitution occurs in any regular manner or is confined to any particular planes in the crystal. Rosenhain* has suggested that the presence of a stranger atom necessitates a distortion of the crystal lattice in its immediate neighbourhood, similar to the effects of mechanical strain, and that the change in properties due to alloying are easily explained on this hypothesis. The greater resistance to deformation can be attributed to irregularities in the slip-planes due to the presence of the larger atoms and consequent distortion in their neighbourhood so that the planes move over each other with greater difficulty. When two of these irregularities meet, they tend to hold up slip altogether. Some such explanation has been applied to the effects of both alloying and mechanical deformation in producing a greater resistance to further distortion.

As far as the writer is aware, the only systematic investigation of the effects of alloying on distortion is that undertaken by Rosbaud and Schmid.† They made zinc crystals by Czochralski's method, containing up to 1·03 per cent. cadmium and 2 per cent. tin. They found that wires made in this way contained the same cadmium and tin content as the molten metal from which they were drawn. Crystals containing 0·53 per cent. cadmium showed no trace of eutectic, but those containing 1·03 per cent. showed a little. On the other hand, 0·1 per cent. tin in zinc forms a eutectic. The distortion was of the same type as pure zinc.

Their experiments showed that the shear stress at the yield point of the alloys of zinc and cadmium was nine times as great with 0·53 per cent. cadmium, and 12·2 times as great with 1·03 per cent. cadmium, as in the case of Kahlbaum zinc (containing 0·03 per cent. cadmium). In the tin-zinc alloys the eutectic

* 'J. Inst. Metals,' vol. 30, p. 5 (1923).

† 'Z. f. Physik,' vol. 32, p. 197 (1925).

separated in layers parallel to the hexagonal basal planes. The shear stress at the yield point for 0·5 per cent. eutectic content was only three times as great as the pure zinc, and this value only increased slightly with increase of eutectic. They also found that the mixed crystals of zinc and cadmium hardened less for the same amount of shear than the purer metal.

These results are in general agreement with those described in the paper on the aluminium-zinc alloys. The actual load at the yield point was not always measured, but the curves in fig. VI, Part I, indicate a rapid rise in shear stress with increasing percentage of zinc. As the extension proceeds the increase in shear stress for the same amount of shear does not differ to anything like the same extent. This seems to indicate that the zinc atoms affect the initial yielding of the metal to a greater extent than the hardening process due to deformation. On the other hand, more than 10 per cent. of zinc reduces the amount of distortion before fracture very considerably. It may be that the amount of shear possible on the slip-planes is less, due to the irregularity of the planes, or it may be that the number of possible slip-planes is reduced. It has certainly been noticed that the slip-bands in these alloys are farther apart than in the pure metal. It is not possible to offer any explanation of the fracture. It generally seems to be associated with slip on a second plane, and it has been shown that the conditions for uniform distortion are more easily maintained when slip is on one plane than when a second plane is involved. There is insufficient evidence, however, on this point.

Turning now to the results of the experiments on brass crystals, they seem to be directly opposite, in many respects, to those of the aluminium-zinc crystals. In the first place, the amount of elongation before fracture is very much greater than for pure copper, and the shear stress at the beginning of the test is very little different from copper. It is possible that the differences observed may be due to the two ways of making the crystals. Crystals of pure metals made by the two methods do not differ, but with alloys that possibility must not be overlooked. Brass crystals cannot be made by the straining method owing to the readiness of this metal to form twin crystals, but it should be possible to make the aluminium-zinc alloys by the melting method, and a final decision on this point awaits an investigation of crystals prepared in this way.*

It has already been pointed out that there are insufficient grounds for thinking that the peculiarities of the distortion of these crystals are due to the direction of growth of the crystals during the process of solidification. But it is possible

* Up to the present it has not been possible to make large aluminium-zinc alloy crystals in this way.

that the alloying atoms take up different positions when the crystals grow from the liquid from what they do when the metal recrystallises in the solid. Further, the part played by mechanical deformation in aiding diffusion is well known, so that it is probable that alloy crystals grown in the solid by a process of straining and heat-treatment are more uniform in structure than those grown from the liquid. One thing, however, is common to both these alloy crystals; the nature of the slip-plane is very much altered by the addition of zinc. The fact that the distortion is unaltered does not prove that the alloy atoms are not concentrated on certain crystal-planes or directions, but X-ray evidence in the case of other alloys is against this view. It is very difficult to account for the fact that in one case the presence of zinc hinders slip, and in the other facilitates slip, if the atoms are evenly distributed and their action is purely mechanical. It appears, therefore, that there must be some other factors involved in the formation of alloys to account for their variety of properties.

I should like to express my thanks to Prof. H. C. H. Carpenter, F.R.S., in whose laboratory this work has been carried out, for his continued interest and help.

I am indebted to Prof. G. I. Taylor, F.R.S., for assistance and advice in the calculations, and to Prof. W. E. Dalby, F.R.S., for his kindness in putting a testing machine in the City and Guilds Engineering College at my disposal.

Meteorological Perturbations of Tides and Currents in an Unlimited Channel rotating with the Earth.

By H. Horrocks, B.A., St. Catharine's College, Cambridge, Isaac Newton Student.

(Communicated by Sir Joseph Larmor, F.R.S.—Received December 9, 1926.)

§ 1. *Introduction.*

The following paper is an extension, to include the Earth's rotation, of a discussion by J. Proudman and A. T. Doodson,* treating of the corresponding phenomena in a non-rotating tidal channel. In the course of the solution it is necessary to have two functions expanded in two related series of special form (§ 9), and a separate paper has been devoted to the expansion-theorem required.†

Equations of Motion.

Let the axis of z be taken vertically upwards, $0x$ transverse to the channel, and $0y$ in the direction of its unlimited extent. Let the axes be so taken that ω, the rotation about the vertical, is positive. Suppose the channel is bounded by $x = 0$, $x = l$, and $z = -h$, and let the mean free surface be $z = 0$. For simplicity we shall assume that the kinematic coefficient of turbulence ν is a function of the depth only.

If we consider disturbances generated by wind and atmospheric pressure varying only with x and t, the variation of the currents and surface elevation will be independent of y. Let u and v (functions of x, z, t) be the perturbations of current components parallel to $0x$ and $0y$, and ζ (x, t) be the elevation of the surface, and p the part of the fluid pressure due to the causes under discussion.

We take as equations of motion, neglecting vertical motion,

$$\left.\begin{aligned} \frac{Du}{Dt} - \nu\frac{\partial^2 u}{\partial x^2} - \frac{\partial}{\partial z}\left(\nu\frac{\partial u}{\partial z}\right) - 2\omega v &= -\frac{1}{\rho}\frac{\partial p}{\partial x} \\ \frac{Dv}{Dt} - \nu\frac{\partial^2 v}{\partial x^2} - \frac{\partial}{\partial z}\left(\nu\frac{\partial v}{\partial z}\right) + 2\omega u &= -\frac{1}{\rho}\frac{\partial p}{\partial y} \\ 0 &= -\frac{1}{\rho}\frac{\partial p}{\partial z} \end{aligned}\right\}. \qquad (1A)$$

* "Time Relations in Meteorological Effects on the Sea," 'Proc. Lond. Math. Soc.,' Series II, vol. 24, part 2.

† *V. infra*, p. 184.

These equations are of the same form as for the problem with a viscous liquid, and differ from those proposed by G. I. Taylor* in not having ν everywhere outside the differential operators. The above form, however, is always consistent with the assumption that the action of the water above a given plane $z = \text{const.}$ on that below it, is equal and opposite to the action of the water below on that above. The form proposed by Taylor and the above form become identical when ν is constant throughout the whole depth.

§ 2. *Approximate Equations of Motion.*

If $\mathfrak{p}$ is the disturbing atmospheric pressure (a function of x, t only) and ζ is the elevation of the surface (also a function of x, t only), the last of the equations (1A) gives

$$p = \mathfrak{p} + \rho g \zeta.$$

Thence the remaining equations give

$$\frac{Du}{Dt} - \nu \frac{\partial^2 u}{\partial x^2} - \frac{\partial}{\partial z}\left(\nu \frac{\partial u}{\partial z}\right) - 2\omega v = -\frac{1}{\rho}\frac{\partial \mathfrak{p}}{\partial x} - g\frac{\partial \zeta}{\partial x}$$

$$\frac{Dv}{Dt} - \nu \frac{\partial^2 v}{\partial x^2} - \frac{\partial}{\partial z}\left(\nu \frac{\partial v}{\partial z}\right) + 2\omega u = 0.$$

Under what circumstances can $\nu \,.\, \partial^2 u/\partial x^2$ and $\nu \,.\, \partial^2 v/\partial x^2$ be neglected? In the problem of steady motion, equations (5) of this discussion would have to be written with $(2i\omega + \kappa^2\nu)\ U_1$ for $2i\omega U_1$ and $(2i\omega - \kappa^2\nu)\ U_2$ for $2i\omega U_2$, if these terms could not be neglected. If L is the width of the channel in kilometres and ϕ the latitude, $\frac{1}{2}\kappa^2\nu/\omega = 216 \,.\, 10^{-8}\,\pi\nu\ (m^2/L^2)$ cosec ϕ. The coefficient of turbulence ν is generally taken to be of the order of 100. Hence for a channel of a few kilometres width or greater, in not too low latitudes, we may neglect these terms provided the m introduced in § 4 is not too large (say < 8). This will include most of the cases of interest.

In the case of free decay (§ 8) we assume a factor $e^{-\lambda \bar{\nu} t/h^2}$ in u and v, where $\bar{\nu}$ is the mean of ν with respect to depth. If $\nu\,\partial^2 u/\partial x^2$ and $\nu\,\partial^2 v/\partial x^2$ could not be neglected, λ in the equations (15) would have to be replaced by $\lambda - (\nu/\bar{\nu})\,h^2\kappa^2$. Now $h^2\kappa^2 = (m\pi h l^{-1})^2$. Thus for a discussion of "long" waves (in our case they are "stationary") this correction can be neglected.

If further we consider small motions (ζ always small compared with h), we may neglect the product terms in our equations of motion, which then become

$$\left.\begin{aligned} \frac{\partial u}{\partial t} - 2\omega v &= \frac{\partial}{\partial z}\left(\nu \frac{\partial u}{\partial z}\right) - g\frac{\partial}{\partial x}(\zeta + \bar{\zeta}) \\ \frac{\partial v}{\partial t} + 2\omega u &= \frac{\partial}{\partial z}\left(\nu \frac{\partial v}{\partial z}\right) \end{aligned}\right\}, \qquad (1)$$

* "Eddy Motion in the Atmosphere," 'Phil. Trans.,' A, vol. 215 (1915).

where $\mathfrak{p} = \rho g\bar{\zeta}$ = pressure due to a layer of liquid of depth $\bar{\zeta}$, and this we shall treat as small.

The equation of continuity is

$$\frac{\partial}{\partial x}\left[\int_{-h}^{\zeta} u \,.\, dz\right] + \frac{\partial \zeta}{\partial t} = 0,$$

neglecting $\zeta\,(\partial u/\partial x)_{z=0}$ and $(d\zeta/\delta x)\,(u)_{z=0}$ this equation becomes

$$\int_{-h}^{0} \frac{\partial u}{\partial x} dz + \frac{\partial \zeta}{\partial t} = 0. \tag{2}$$

§ 3. *Boundary Conditions.*

It is evident that $u = 0$ when $x = 0$ or l. We shall assume further that there is no slipping at the boundaries other than the free surface, *i.e.*, $u = v = 0$ when $x = 0$ or l and when $z = -h$.

At the free surface $\partial u/\partial z$ and $\partial v/\partial z$ are proportional to the tangential force components, parallel to $0x$ and $0y$, due to the wind. The resultant tangential force depends on the velocity of the wind, and is in the same direction. The wind velocity at the water surface, however, is not in the direction of the downward pressure gradient, but is inclined to this direction at an angle varying from 40° to 90°, to the right in the northern hemisphere and to the left in the southern. Thus our remaining boundary conditions can be expressed in the form

$$\left(\frac{\nu}{\bar{\nu}}\right)_0 \left(\frac{\partial u}{\partial z}\right)_{z=0} = -X\frac{\partial \mathfrak{p}}{\partial x}, \qquad \left(\frac{\nu}{\bar{\nu}}\right)_0 \left(\frac{\partial v}{\partial z}\right)_{z=0} = Y\frac{\partial \mathfrak{p}}{\partial x},$$

where $\bar{\nu}$ = the mean value of ν in the range considered, and X and Y are functions of x and t, both being positive, whether we consider the northern or southern hemisphere.

§ 4. *Subsidiary Equations.*

Suppose x enters into $\bar{\zeta}$ only through a factor $\cos \kappa x$. We substitute the following forms for $\bar{\zeta}$, ζ, u, v :—

$$\left.\begin{aligned} \kappa\bar{\zeta} &= P\cos\kappa x, & (\bar{\nu}/gh^2)\,u &= U\sin\kappa x \\ \kappa\zeta &= A\cos\kappa x, & (\bar{\nu}/gh^2)\,v &= V\sin\kappa x \end{aligned}\right\}. \tag{3}$$

These forms satisfy the boundary conditions at $x = 0$ and $x = l$ if $\kappa l = m\pi$, where m is any integer, which we may suppose positive ; P and A being functions of t and U and V being functions of z. If we put also

$$U_1 = U + iV, \qquad U_2 = U - iV, \tag{4}$$

the remaining conditions to be satisfied are

$$\left.\begin{aligned}\frac{\partial U_1}{\partial t} + 2i\omega U_1 &= \frac{\bar{\nu}}{h^2}\left\{\frac{\partial}{\partial \xi}\left(\eta \frac{\partial U_1}{\partial \xi}\right)\right\} + \frac{\bar{\nu}}{h^2}(A+P)\\ \frac{\partial U_2}{\partial t} - 2i\omega U_2 &= \frac{\bar{\nu}}{h^2}\left\{\frac{\partial}{\partial \xi}\left(\eta \frac{\partial U_2}{\partial \xi}\right)\right\} + \frac{\bar{\nu}}{h^2}(A+P)\end{aligned}\right\} \tag{5}$$

$$\int_0^1 (U_1 + U_2)\, d\xi = \frac{-2\bar{\nu}}{\kappa^2 g h^3}\frac{dA}{dt}, \tag{6}$$

$$\left.\begin{aligned}&U_1 = U_2 = 0 \quad \text{when} \quad \xi = 1,\\ &\frac{\partial U_1}{\partial \xi} = -We^{i\beta}; \quad \frac{\partial U_2}{\partial \xi} = -We^{-i\beta}; \quad \text{when } \xi = 0,\end{aligned}\right\} \tag{7}$$

where $\eta = (\nu/\bar{\nu})$ is a function of z; $\xi = -z/h$; and W, β are supposed to be independent of x in the present development. We shall also use the notation $We^{i\beta} = W_x + iW_y$.

§ 5. *Steady State.*

Suppose the steady state is established in which $\bar{\zeta}$, ζ, u, v, W_x, W_y are independent of t, U and V, depending on z only.

If in this case we put

$$U_1 = (A+P)\,G(\xi) \qquad U_2 = (A+P)\,F(\xi) \qquad W = (A+P)\,\overline{W},$$

in (5), (6), (7) we have, putting $\tilde{\omega}^2 = \omega h^2/\bar{\nu}$,

$$\frac{\partial}{\partial \xi}\left[\eta(\xi)\frac{\partial G(\xi)}{\partial \xi}\right] - 2i\tilde{\omega}^2 G(\xi) = -1$$

$$\frac{\partial}{\partial \xi}\left[\eta(\xi)\frac{\partial F(\xi)}{\partial \xi}\right] + 2i\tilde{\omega}^2 F(\xi) = -1$$

$$\int_0^1 [F(\xi) + G(\xi)]\, d\xi = 0,$$

whilst equations (7) remain of similar form. Evidently the variation of the current with depth, determined by $F(\xi)$ and $G(\xi)$, will be the same in all cases when the *ratio* of the components of tangential force and the total atmospheric pressure is unchanged, *i.e.*, when $\overline{W}$ and β are unchanged. In the present case of course $F(\xi)$ and $G(\xi)$ are conjugate, and it is sufficient to consider only one of them to find the motion.

Considering in particular the case when the coefficient of turbulence is constant throughout the whole depth and putting $j = 1 + i$

$$G(\xi) = c_1 e^{j\varpi\xi + i\epsilon_1} + c_2 e^{-j\varpi\xi + i\epsilon_2} - \tfrac{1}{2}i/\tilde{\omega}^2, \tag{8}$$

where c_1, c_2, ε_1, ε_2 are real and must be determined by conditions (7), and then A may be found from the equation of continuity.

We require

$$c_1 e^{j\varpi + i\epsilon_1} + c_2 e^{-j\varpi + i\epsilon_2} = \tfrac{1}{2} i/\tilde{\omega}^2$$

$$c_1 e^{i\epsilon_1} - c_2 e^{i\epsilon_2} = (-1/j\tilde{\omega})\,\overline{W} e^{i\beta}.$$

Hence

$$c_1 e^{i\epsilon_1} = \tfrac{1}{2}\,\mathrm{sech}\, j\tilde{\omega}\left[\tfrac{1}{2} i/\tilde{\omega}^2 - (1/\tilde{\omega}\sqrt{2})\,\overline{W} e^{i\left(\beta - \frac{\pi}{4}\right) - j\varpi}\right]$$

$$c_2 e^{i\epsilon_2} = \tfrac{1}{2}\,\mathrm{sech}\, j\tilde{\omega}\left[\tfrac{1}{2} i/\tilde{\omega}^2 + (1/\tilde{\omega}\sqrt{2})\,\overline{W} e^{i\left(\beta - \frac{\pi}{4}\right) + j\varpi}\right].$$

Hence

$$U_1 = (A + P)\left\{\mathrm{sech}\, j\tilde{\omega}\left[(\tfrac{1}{2} i/\tilde{\omega}^2)\cosh j\tilde{\omega}\xi + (1/\tilde{\omega}\sqrt{2})\,\overline{W} e^{i\left(\beta - \frac{\pi}{4}\right)} \sinh j\tilde{\omega}(1-\xi)\right] - \tfrac{1}{2} i/\tilde{\omega}^2\right\}. \quad (10)$$

The equation of continuity to determine A in terms of P and W is in this case

$$\Re\left\{(1/j\tilde{\omega})\left[c_1 e^{i\epsilon_1}(e^{j\varpi} - 1) - c_2 e^{i\epsilon_2}(e^{-j\varpi} - 1)\right] - (\tfrac{1}{2} i/\tilde{\omega}^2)\right\} = 0,$$

i.e.,

$$\Re\left\{j(A + P)\tanh j\tilde{\omega} + 2i\tilde{\omega} W e^{i\beta}(\mathrm{sech}\, j\tilde{\omega} - 1)\right\} = 0. \quad (11)$$

§ 6. *Case when $\tilde{\omega}$ is Small.*

Taking terms of lowest order in $\tilde{\omega}$, for equation (11) we find

$$\tfrac{2}{3}(A + P) + W_x = 0.$$

Approximating also in (10) we find

$$U_1 = \tfrac{3}{4} W_x(\xi^2 - 1) + (W_x + iW_y)(1 - \xi).$$

Hence in this case, corresponding to a pressure distribution given by $\kappa\overline{\zeta} = P\cos\kappa x$, and our tangential force (7) we have

$$\left.\begin{aligned} \kappa\zeta &= -(P + \tfrac{3}{2} W_x)\cos\kappa x \\ (\overline{\nu}/gh^2)\,u &= \tfrac{3}{4} W_x(z/h + 1)(z/h + \tfrac{1}{3})\sin\kappa x \\ (\overline{\nu}/gh^2)\,v &= W_y(z/h + 1)\sin\kappa x \end{aligned}\right\}. \quad (12)$$

If $W_y = 0$ as in the case of no rotation, the above solution takes the same form as the corresponding case in the problem which Proudman and Doodson have discussed.

On examining the order of the value of $\tilde{\omega}$ required for our approximations to hold, we find that for rates of rotation comparable with those found at

points on the Earth, the channel would have to be too shallow to be of interest, except within a few kilometres of the equator.

§ 7. *Case when $\tilde{\omega}$ is Large.*

Approximating in equation (11) we find on retaining the most important terms in the two expressions

$$A + P = 2\tilde{\omega} W_y.$$

Hence

$$\kappa\zeta = -(P - 2\tilde{\omega} W_y) \cos \kappa x. \tag{13}$$

For the rates of rotation found in latitudes higher than 30°, this approximation would require the depth of the channel to be about 50 metres or more.

Near surface

$$U_1 = (1/\tilde{\omega}\sqrt{2})\ We^{-\varpi\xi} e^{i\left(\beta - \frac{\pi}{4} - \varpi\xi\right)} - (i/2\tilde{\omega}^2)(A + P).$$

Near base (where $\xi' = 1 - \xi$)

$$U_1 = (i/2\tilde{\omega}^2)(A + P)(e^{-j\varpi\xi'} - 1).$$

In the middle

$$U_1 = -(i/2\tilde{\omega}^2)(A + P) \quad \text{approximately.*}$$

§ 8. *Free Variable State.*

We now investigate the motion when there is no meteorological disturbance, but an initial current-system and surface elevation decay through the effects of turbulence.

Instead of equations (3) we consider the possibility of solutions of the form

$$\left.\begin{aligned} \kappa\zeta &= A_\lambda e^{-\lambda\bar{\nu}t/h^2} \cos \kappa x \\ U_1 &= A_\lambda \psi_\lambda(\xi)\, e^{-\lambda\bar{\nu}t/h^2} \\ U_2 &= A_\lambda \phi_\lambda(\xi)\, e^{-\lambda\bar{\nu}t/h^2} \end{aligned}\right\}, \tag{14}$$

where U_1 and U_2 are defined as in (4) and A_λ is constant.

The equations (1) then give rise to equations

$$\left.\begin{aligned} \frac{\partial}{\partial\xi}\left[\eta(\xi)\frac{\partial\phi_\lambda(\xi)}{\partial\xi}\right] + (\lambda + i\gamma)\,\phi_\lambda(\xi) &= -1 \\ \frac{\partial}{\partial\xi}\left[\eta(\xi)\frac{\partial\psi_\lambda(\xi)}{\partial\xi}\right] + (\lambda - i\gamma)\,\psi_\lambda(\xi) &= -1 \end{aligned}\right\}, \tag{15}$$

where $\gamma = 2\tilde{\omega}^2 = 2\omega h^2/\bar{\nu}$, which is positive.

* W. Ekman (' Arkiv för Matematik,' vol. 2 (1905–6), "On the Influence of the Earth's Rotation on Ocean Currents") in a paper on currents generated by a uniform wind on an unbounded ocean, gives diagrams illustrating the variation of current with depth, and these apply also in the present case.

The boundary conditions for the free state are

$$\phi_\lambda(1) = \psi_\lambda(1) = 0, \quad \phi_\lambda'(0) = \psi_\lambda'(0) = 0, \tag{16}$$

and the equation of continuity, for determining suitable values of λ, is

$$\tfrac{1}{2}\int_0^1 [\phi_\lambda(\xi) + \psi_\lambda(\xi)]\, d\xi = \alpha\lambda, \tag{17}$$

where $\alpha = (\bar{\nu})^2 g^{-1}\kappa^{-2}h^{-5}$, which is positive.

The properties of the solutions of these differential equations which satisfy the required conditions have been discussed in connection with the expansion theorem referred to in § 1. The expansions necessary to deal with the decay of any initial disturbance of the form

$$u = \bar{U}(z)\sin\kappa x, \qquad v = \bar{V}(z)\sin\kappa x, \qquad \zeta = \mathrm{K}\cos\kappa x$$

have also, as we shall see, been established.

We shall discuss in particular the case when the coefficient of turbulence is constant throughout the water [$\eta(\xi) = 1$], in which case the possible solutions are

$$\left.\begin{aligned} \phi_\lambda(\xi) &= \frac{1}{\lambda + i\gamma}\left[\frac{\cos(\lambda + i\gamma)^{\frac{1}{2}}\xi}{\cos(\lambda + i\gamma)^{\frac{1}{2}}} - 1\right] \\ \psi_\lambda(\xi) &= \frac{1}{\lambda - i\gamma}\left[\frac{\cos(\lambda - i\gamma)^{\frac{1}{2}}\xi}{\cos(\lambda - i\gamma)^{\frac{1}{2}}} - 1\right] \end{aligned}\right\}, \tag{18}$$

where λ satisfies

$$\alpha\lambda = \mathrm{N}(\lambda) = \tfrac{1}{2}\left[\frac{1}{(\lambda + i\gamma)^{\frac{3}{2}}}\tan(\lambda + i\gamma)^{\frac{1}{2}} + \frac{1}{(\lambda - i\gamma)^{\frac{3}{2}}}\tan(\lambda - i\gamma)^{\frac{1}{2}} - \frac{2\lambda}{\lambda^2 + \gamma^2}\right]. \tag{19}$$

We know the nature of the distribution of the roots of this equation, and thence we see that there is an odd (finite) number of possible solutions in which the motion is one of non-oscillatory decay, and an infinity of possible cases of oscillatory decay. In the case of oscillatory decay, λ being complex, the solutions must be grouped in the form

$$\begin{aligned} \kappa\zeta &= A_\lambda[e^{-\lambda\bar{\nu}t/h^2} + e^{-\bar{\lambda}\bar{\nu}t/h^2}]\cos\kappa x \\ U_1 &= A_\lambda[\psi_\lambda(\xi)\, e^{-\lambda\bar{\nu}t/h^2} + \psi_{\bar{\lambda}}(\xi)\, e^{-\lambda\bar{\nu}t/h^2}] \\ U_2 &= A_\lambda[\phi_\lambda(\xi)\, e^{-\lambda\bar{\nu}t/h^2} + \phi_{\bar{\lambda}}(\xi)\, e^{-\bar{\lambda}\bar{\nu}t/h^2}] \end{aligned}$$

or

$$\kappa\zeta = -iA_\lambda[e^{-\lambda\bar{\nu}t/h^2} - e^{-\bar{\lambda}\bar{\nu}t/h^2}]\cos\kappa x, \text{ etc.},$$

where λ and $\bar{\lambda}$ are conjugate roots of $\mathrm{N}_1'(\lambda) = \alpha\lambda$.

§ 9. *Decay of a Given Disturbance.*

Suppose there is an initial surface elevation given by

$$\kappa\zeta = A\cos\kappa x$$

and a velocity distribution given by

$$U_1 = Ag_1(\xi) \qquad U_2 = Af_1(\xi) \Bigg\}. \tag{20}$$

We have seen in the expansion theorem, of which we here use the notation, how $f_1(\xi)$ and $g_1(\xi)$ can be expanded in the form

$$f_1(\xi) = \sum_{n=1}^{\infty} C_n\phi_n(\xi) \qquad g_1(\xi) = \sum_{n=1}^{\infty} C_n\psi_n(\xi),$$

where λ_n satisfies the equation $\mathrm{N}(\lambda) = \alpha\lambda$, and $\sum_{n=1}^{\infty} C_n = 1$.

If we now put

$$\left.\begin{aligned} \kappa\zeta &= A\cos\kappa x \sum_{n=1}^{\infty} C_n e^{-\lambda_n\bar{\nu}t/h^2} \\ U_1 &= A\sum_{n=1}^{\infty} C_n\psi_n(\xi)\, e^{-\lambda_n\bar{\nu}t/h^2} \\ U_2 &= A\sum_{n=1}^{\infty} C_n\phi_n(\xi)\, e^{-\lambda_n\bar{\nu}t/h^2} \end{aligned}\right\}, \tag{21}$$

we have the initial forms we require, and our differential equations and boundary conditions are satisfied, and also the equation of continuity.

It is easy to see that the above expression for ζ remains real for all values of t, for

$$C_n = \frac{-\alpha + \frac{1}{2}\int_0^1 [f_1(\xi)\,\phi_n(\xi) + g_1(\xi)\,\psi_n(\xi)]\,d\xi}{\mathrm{N}'(\lambda_n) - \alpha}. \tag{22}$$

Since $f_1(\xi)$ and $g_1(\xi)$ are conjugate and $\psi_n(\xi)$ and $\phi_n(\xi)$ are conjugate when λ_n, $\lambda_{\bar{n}}$ are, we see at once that C_n, $C_{\bar{n}}$ are conjugate. Whence $C_n e^{-\lambda_n\bar{\nu}t/h^2} + C_n e^{-\lambda_{\bar{n}}\bar{\nu}t/h^2}$ is real, and so is the above expression for ζ. Similarly $u \propto (U_1 + U_2)$ and $v \propto (U_1 - U_2)$ are real.

Of the above expressions, that for ζ can be expressed as an integral. Consider the function

$$\frac{-\alpha + \frac{1}{2}\int_0^1 [f_1(\xi)\,\phi_\lambda(\xi) + g_1(\xi)\,\psi_\lambda(\xi)]\,d\xi}{\mathrm{N}(\lambda) - \alpha\lambda}\, e^{-\lambda\bar{\nu}t/h^2}.$$

As in the expansion discussion, if this is integrated round a contour consisting of the semi-circle of radius $r^2\pi^2$, and the imaginary axis between $ir^2\pi^2$, and

$-ir^2\pi^2$, the contribution due to the semi-circle tends to zero as r tends to infinity. The poles of this function are the zeros of $N(\lambda) - \alpha\lambda$, and since these have all got real parts positive in our case ($\alpha > 0$) it follows that

$$\int_{+i\infty}^{-i\infty} \frac{-\alpha + \frac{1}{2}\int_0^1 [f_1(\xi)\,\phi_\lambda(\xi) + g_1(\xi)\,\psi_\lambda(\xi)]\,d\xi}{N(\lambda) - \alpha\lambda} e^{-\lambda\bar{\nu}t/h^2}\,d\lambda = 2\pi i \sum_{n=1}^{\infty} C_n e^{-\lambda_n \bar{\nu} t/h^2}.$$

Hence

$$\kappa\zeta = \frac{A\cos\kappa x}{2\pi i}\int_{-i\infty}^{+i\infty} \frac{\alpha - \frac{1}{2}\int_0^1 [f_1(\xi)\,\phi_\lambda(\xi) + g_1(\xi)\,\psi_\lambda(\xi)]\,d\xi}{N(\lambda) - \alpha\lambda} e^{-\lambda\bar{\nu}t/h^2}\,d\lambda. \qquad (23)$$

We may notice here that if in (20) we had expressed the velocity terms in the form $U_1 = g_1(\xi)$, $U_2 = f_1(\xi)$, the factor A would not appear in the expressions for ζ, U_1, U_2, in (21), whilst instead of (22) we should have

$$C_n = \frac{-A\alpha + \frac{1}{2}\int_0^1 [f_1(\xi)\,\phi_n(\xi) + g_1(\xi)\,\psi_n(\xi)]\,d\xi}{N'(\lambda_n) - \alpha}.$$

A corresponding change would be introduced into (23).

§ 10. *Variable State with Constant Meteorological Conditions.*

Noting that (5) (6) and the first pair of conditions in (7) are homogeneous and of the first degree in A, P, U_1, U_2, it follows that any linear combination of solutions of these equations is also a solution. From the second pair in (7) we see that in such a combination the wind-effect required is the corresponding linear combination of the original wind-effects. In the present case we can take the sum of two solutions, one being the steady state corresponding to the given pressure distribution and wind, the other being a suitable case of free decay.

Suppose the initial state is specified by

$$\left.\begin{aligned} \kappa\zeta &= A\cos\kappa x \\ U_1 &= g_1(\xi) \qquad\qquad U_2 = f_1(\xi) \end{aligned}\right\} \qquad (24)$$

and the meteorological conditions maintained constant from time $t = 0$ onwards are

$$\left.\begin{aligned} \kappa\bar{\zeta} &= P\cos\kappa x \\ \left[\frac{\partial U_1}{\partial \xi}\right]_0 &= -We^{i\beta} \qquad \left[\frac{\partial U_2}{\partial \xi}\right]_0 = -We^{-i\beta} \end{aligned}\right\}. \qquad (25)$$

We have seen how to find the corresponding steady state, and we shall suppose that this is given by

$$\left.\begin{aligned} \kappa\zeta &= B\cos\kappa x \\ U_1 = g(\xi) &= (B+P)\,G(\xi), \quad U_2 = f(\xi) = (B+P)\,F(\xi) \end{aligned}\right\}, \tag{26}$$

where B, $g(\xi)$, $f(\xi)$, involve P, W, and β in a form that can be determined.

Thus to (26) we have to add the solution corresponding to the free decay of the conditions

$$\kappa\zeta = (A-B)\cos\kappa x$$

$$U_1 = g_1(\xi) - g(\xi) = \bar{g}(\xi) \text{ say}$$

$$U_2 = f_1(\xi) - f(\xi) = \bar{f}(\xi) \text{ say}.$$

Hence the solution for the establishment of a steady state is

$$\left.\begin{aligned} \kappa\zeta &= \cos\kappa x \sum_{n=1}^{\infty} C_n e^{-\lambda_n \bar{\nu} t/h^2} + B\cos\kappa x \\ U_1 &= \sum_{n=1}^{\infty} C_n \psi_n(\xi)\, e^{-\lambda_n \bar{\nu} t/h^2} + g(\xi) \\ U_2 &= \sum_{n=1}^{\infty} C_n \phi_n(\xi)\, e^{-\lambda_n \bar{\nu} t/h^2} + f(\xi) \end{aligned}\right\}, \tag{27}$$

where

$$C_n = \frac{(B-A)\,\alpha + \frac{1}{2}\int_0^1 [\bar{f}(\xi)\,\phi_n(\xi) + \bar{g}(\xi)\,\psi_n(\xi)]\,d\xi}{N'(\lambda_n) - \alpha}.$$

§ 11. If we suppose the water to have been at rest with its surface horizontal when the wind and pressure system was set up, we should have

$$C_n = \frac{B\alpha - \frac{1}{2}\int_0^1 [f(\xi)\,\phi_n(\xi) + g(\xi)\,\psi_n(\xi)]\,d\xi}{N'(\lambda_n) - \alpha}. \tag{28}$$

This expression can be thrown into a form depending more directly on the wind and pressure system, thus :—

$$(\lambda + i\gamma)\int_0^1 F\phi_\lambda\,d\xi = \int_0^1 F\left[-1 - \frac{\partial}{\partial\xi}\left(\eta\frac{\partial\phi_\lambda}{\partial\xi}\right)\right]d\xi$$

$$= -\int_0^1 F\,d\xi - \left[F\eta\frac{\partial\phi_\lambda}{\partial\xi}\right]_0^1 + \int_0^1 F'\eta\frac{\partial\phi_\lambda}{\partial\xi}\,d\xi.$$

Thus

$$(\lambda_n + i\gamma)\int_0^1 F\phi_n\,d\xi = -\int_0^1 F\,d\xi + \int_0^1 \eta F'\phi_n'\,d\xi$$

and

$$(\lambda_n - i\gamma)\int_0^1 G\psi_n\,d\xi = -\int_0^1 G\,d\xi + \int_0^1 \eta G'\psi_n'\,d\xi,$$

similarly

$$i\gamma\int_0^1 F\phi_n\,d\xi = -\int_0^1 \phi_n\,d\xi + \int_0^1 \eta F'\phi_n'\,d\xi + \eta(0)\,\phi_n(0)\,F'(0)$$

and

$$-i\gamma\int_0^1 G\psi_n\,d\xi = -\int_0^1 \psi_n\,d\xi + \int_0^1 \eta G'\psi_n'\,d\xi + \eta(0)\,\psi_n(0)\,G'(0).$$

Hence

$$\lambda_n\int_0^1 (F\phi_n + G\psi_n)\,d\xi = \int_0^1 (\phi_n + \psi_n)\,d\xi - \int_0^1 (F+G)\,d\xi - \eta_0(\phi_n F' + G'\psi_n)_0,$$

i.e.,

$$\int_0^1 (F\phi_n + G\psi_n)\,d\xi = 2\alpha + \eta_0\overline{W}\,[\phi_n(0)\,e^{-i\beta} + \psi_n(0)\,e^{i\beta}]\,(1/\lambda_n).$$

To find C_n the above must be multiplied by $\frac{1}{2}(B+P)$ and subtracted from $B\alpha$ and the whole divided by $N'(\lambda_n) - \alpha$. We thus obtain

$$C_n = \frac{-\alpha P - \frac{1}{2}\eta(0)\,W[\phi_n(0)\,e^{-i\beta} + \psi_n(0)\,e^{i\beta}]\,(1/\lambda_n)}{N'(\lambda_n) - \alpha}.$$

This expresses C_n in terms of λ_n and the wind and pressure conditions. The surface elevation ζ could be expressed again partly in terms of a contour integral.

§ 12. *Expressions for the Surface Elevation and Currents in the General Variable State.*

Suppose that in a short interval of time $\delta\tau = \tau_n - \tau_{n-1}$ the conditions (25) are slightly varied, P, W, β being functions of τ ($<t$)

$$\left.\begin{aligned} \delta(\kappa\bar{\zeta}) &= \dot{P}(\tau)\,\delta\tau\cos\kappa x \\ \delta\left[\frac{\partial U_1}{\partial\xi}\right]_0 = -\frac{d}{d\tau}[We^{i\beta}]\,\delta\tau; &\qquad \delta\left[\frac{\partial U_2}{\partial\xi}\right]_0 = -\frac{d}{d\tau}[We^{-i\beta}]\,\delta\tau \end{aligned}\right\}. \tag{29}$$

There will be a corresponding change in (26), which give the steady state corresponding to the wind and pressure conditions which hold at time τ.

$$\left.\begin{aligned} \delta(\kappa\zeta) &= \dot{B}(\tau)\,\delta\tau\cos\kappa x \\ \delta(U_1) = \dot{g}(\xi,\tau)\,\delta\tau; &\qquad \delta(U_2) = \ddot{f}(\xi,\tau)\,\delta\tau \end{aligned}\right\}. \tag{30}$$

Now the change in the motion which would otherwise occur, which arises when the conditions are changed, is the same as the motion which would be set up in water at rest when conditions equal to the change in wind and pressure are applied. Thus if the motion of the water and the amplitude of its surface elevation and conditions (25) are given at time τ_1, the "configuration" of the water at time t will be found by adding the effects of a set of pressure and wind distributions given by (29) acting on water initially at rest through intervals τ_n to t respectively, to the effect of the initial pressure and wind acting unchanged on the initial configuration from time τ_1 to time t.

Hence at time t we have

$$\kappa\zeta = \int_{\tau_1}^{t} \cos\kappa x \left[\sum_{n=1}^{\infty} \dot{C}_n(\tau)\, e^{-\lambda_n \bar{\nu}(t-\tau)/h^2} + \dot{B}(\tau)\right] d\tau + B(\tau_1)\cos\kappa x$$
$$+ \cos\kappa x \sum_{n=1}^{\infty} [C_n(\tau_1) - K_n(\tau_1)]\, e^{-\lambda_n \bar{\nu}(t-\tau_1)/h^2},$$

$$U_1 = \int_{\tau_1}^{t} \left[\sum_{n=1}^{\infty} \dot{C}_n(\tau)\, \psi_n(\xi)\, e^{-\lambda_n \bar{\nu}(t-\tau)/h^2} + \dot{g}(\xi, \tau)\right] d\tau + g(\xi, \tau_1)$$
$$+ \sum_{n=1}^{\infty} [C_n(\tau_1) - K_n(\tau_1)]\, \psi_n(\xi)\, e^{-\lambda_n \bar{\nu}(t-\tau_1)/h^2},$$

with a similar expression for U_2. Here we use

$$C_n(\tau) = \frac{\alpha B(\tau) - \frac{1}{2}\int_0^1 [f(\xi, \tau)\, \phi_n(\xi) + g(\xi, \tau)\, \psi_n(\xi)]\, d\xi}{N'(\lambda_n) - \alpha}$$

and $K_n(\tau)$ is a similar expression with $A(\tau_1)$, $g_1(\xi, \tau_1)$, $f_1(\xi, \tau_1)$ substituted for $B(\tau)$, $g(\xi, \tau)$, $f(\xi, \tau)$ respectively. Integrating the terms directly integrable and letting $\tau_1 \longrightarrow -\infty$

$$\left.\begin{aligned} \kappa\zeta &= \int_{-\infty}^{t} \cos\kappa x \sum_{n=1}^{\infty} \dot{C}_n(\tau)\, e^{-\lambda_n \bar{\nu}(t-\tau)/h^2}\, d\tau + B(t)\cos\kappa x \\ U_1 &= \int_{-\infty}^{t} \sum_{n=1}^{\infty} \dot{C}_n(\tau)\, \psi_n(\xi)\, e^{-\lambda_n \bar{\nu}(t-\tau)/h^2}\, d\tau + g(\xi, t) \\ U_2 &= \int_{-\infty}^{t} \sum_{n=1}^{\infty} \dot{C}_n(\tau)\, \phi_n(\xi)\, e^{-\lambda_n \bar{\nu}(t-\tau)/h^2}\, d\tau + f(\xi, t) \end{aligned}\right\}. \qquad (31)$$

§ 13. *A Particular Case.*

To obtain an indication of the nature of the phenomena in the establishment of the steady state, let us consider the case when a pressure distribution

$$\kappa\bar{\zeta} = -A\cos\kappa x$$

is suddenly applied, and maintained constant in the case when $m = 1$ and

$$\gamma = 2\omega h^2/\bar{\nu} = 100, \qquad \alpha = \bar{\nu}^2/g\kappa^2 h^5 = 10^{-7},$$

corresponding roughly with a channel 100 metres deep and 40 km. across in latitude 55° N., the turbulence coefficient with a value of 80 being considered constant.

We must add the surface elevation $\kappa\zeta = A \cos \kappa x$ to the free decay of $\kappa\zeta = -A \cos \kappa x$, $u = 0$, $v = 0$. This can be done by the use of equation (23), employing an approximate formula for the range in which $|\lambda|$ is large enough to approximate, and mechanical methods of integration for the range of smaller values of $|\lambda|$.

We find that the oscillations of the surface on either side of the configuration to be set up have a period of about 0·19 of a day, and that their amplitude at the end of one day is about 1/4 of the final surface elevation.

§ 14. In their Reports for 1920 and 1921, the Liverpool Tidal Institute published graphs giving the residues when the predicted tide had been removed from the observed tide, extending over several days. These graphs frequently give evidence of oscillations, whose period is between 1/5 and 1/6 of a day, continuing for a time. We reproduce here the curve for June 29, 1918, showing evidence of a decaying oscillation of this period approximately, persisting for four periods and then apparently renewed.

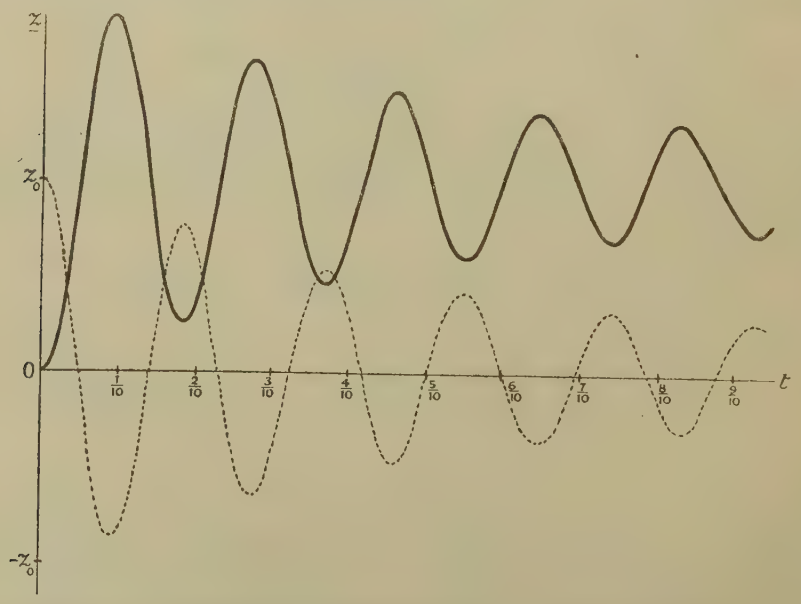


Fig. 1.—Graphs for Establishment (continuous line) and Free Decay (broken line) of Steady State set up by Pressure Distribution.

$\kappa\zeta = -A \cos \kappa x$,

t is time in days; z is surface elevation; z_0 is $(A/\kappa) \cos \kappa x$.

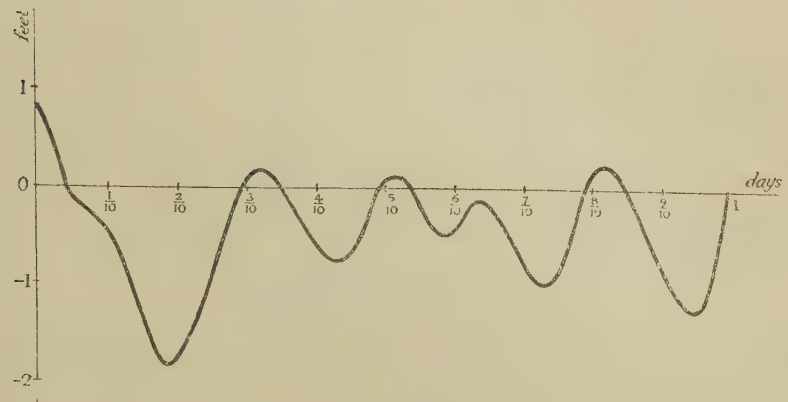


FIG. 2.—Difference between Predictions and Gauge Record at Liverpool on June 29, 1918.

Now the predicted tide is obtained by harmonic analysis of records extending over a long period, and should contain all important periodic terms occurring regularly and derived directly or indirectly from astronomical causes. Irregular features such as meteorological effects would not occur in the predictions, and should be important in the curve of residues.

It may thus be possible to regard these residual oscillations as being due to changes in meteorological conditions over the Irish Sea generating a disturbance in the deep channel running north and south between St. George's Channel and the North Channel, whose dimensions approximate to those in our numerical example, this then being propagated over the shallow eastern waters of the sea.

Generalised Sturm-Liouville Expansions in Series of Pairs of Related Functions.

By H. Horrocks, B.A., St. Catharine's College, Cambridge, Isaac Newton Student.

(Communicated by Sir Joseph Larmor, F.R.S.—Received December 9, 1926.)

§ 1. *Statement of Problem.*

The expansions here developed are required for the author's discussion of "Meteorological Perturbations of Tides and Currents in an Unlimited Channel rotating with the Earth" (*v. supra*, p. 170).

Let $\eta(x)$ be a real differentiable function of x defined in the range $0 \leqslant x \leqslant 1$, and satisfying the condition $\eta(x) > c > 0$ for all such x. Let $\phi_\lambda(x)$ and $\psi_\lambda(x)$ be functions of the real variable x and the complex parameter λ, defined in the above range by the equations

$$\left.\begin{aligned} \frac{d}{dx}\left[\eta(x)\frac{d\phi_\lambda(x)}{dx}\right] + (\lambda + i\gamma)\,\phi_\lambda(x) &= -1 \\ \frac{d}{dx}\left[\eta(x)\frac{d\psi_\lambda(x)}{dx}\right] + (\lambda - i\gamma)\,\psi_\lambda(x) &= -1 \end{aligned}\right\}, \tag{1}$$

together with the boundary conditions

$$\phi'_\lambda(0) = 0, \quad \psi'_\lambda(0) = 0, \quad \phi'_\lambda(1) = 0, \quad \psi_\lambda(1) = 0, \tag{2}$$

γ being a prescribed constant.

Let $f(x)$, $g(x)$ be two functions of x defined in the same range, and let $f_0(x) = g_0(x) = 0$ in this range.

We require to expand $f(x)$ and $g(x)$ in terms of the functions $\phi_\lambda(x)$ and $\psi_\lambda(x)$ respectively, for a suitable sequence of values of λ, the sequences of coefficients occurring in both expansions being identical, and forming a series whose sum is zero. $f_0(x)$ and $g_0(x)$ are to be expanded in a similar form, except that in this case the coefficients are to form a series whose sum is unity.

The form of the series will be found in § 9.

We shall see that this is possible provided $f(x)$ and $g(x)$ have Sturm-Liouville expansions in terms of the functions $V_\mu(x)$ satisfying the conditions (3) to (5).

$$\frac{d}{dx}\left[\eta(x)\frac{dV_\mu(x)}{dx}\right] + \mu V_\mu(x) = 0. \tag{3}$$

$$V_\mu(0) = 1, \qquad V'_\mu(0) = 0, \tag{4}$$

$$V_\mu(1) = 0. \tag{5}$$

The values of μ for which these conditions are satisfied are known to be real, positive, and infinite in number, say $\nu_1, \nu_2, \ldots \nu_n, \ldots$ and $\nu_n \to \infty$ as $n \to \infty$ *. We denote by $V_1(x)$, $V_2(x), \ldots V_n(x), \ldots$ the functions $V_\mu(x)$ corresponding to these values of μ.

§ 2. *Auxiliary Functions.*

Let the solution of equations (3) and (4) for any value of μ be denoted by $v_\mu(x)$ and put

$$\Phi_\lambda(x) = v_{\lambda+i\gamma}(x), \qquad \Psi_\lambda(x) = v_{\lambda-i\gamma}(x). \tag{6}$$

$\phi_\lambda(x)$ and $\psi_\lambda(x)$ can at once be expressed in terms of these auxiliary functions whose properties are known, thus :—

$$\left.\begin{aligned} \phi_\lambda(x) &= \frac{1}{\lambda + i\gamma}\left[\frac{\Phi_\lambda(x)}{\Phi_\lambda(1)} - 1\right] \\ \psi_\lambda(x) &= \frac{1}{\lambda - i\gamma}\left[\frac{\Psi_\lambda(x)}{\Psi_\lambda(1)} - 1\right] \end{aligned}\right\}, \tag{7}$$

whenever these exist. Thus these equations give $\phi_\lambda(x)$ and $\psi_\lambda(x)$ for all values of λ except $-i\gamma$, $\nu_1 - i\gamma, \ldots \nu_n - i\gamma, \ldots$ when $\phi_\lambda(x)$ is not defined by (7), and $+i\gamma$, $\nu_1 + i\gamma, \ldots \nu_n + i\gamma, \ldots$ when $\psi_\lambda(x)$ is not defined by (7).

Returning to the equations (1) and (2) we see that when

when

$$\left.\begin{aligned} \lambda &= -i\gamma \qquad \phi_\lambda(x) = \int_x^1 [x/\eta(x)]\,dx \\ \lambda &= +i\gamma \qquad \psi_\lambda(x) = \int_x^1 [x/\eta(x)]\,dx \end{aligned}\right\} \tag{7A}$$

These exist and are finite for all x in $0 \leqslant x \leqslant 1$.

§ 3. *Characteristic Constants and Functions.*

$\Phi_\lambda(x)$, $\Psi_\lambda(x)$, $\Phi'_\lambda(x)$, $\Psi'_\lambda(x)$, where dashes denote (as throughout this paper) differentiations with respect to x, are integral functions of λ when regarded as functions of this argument. $\Phi_\lambda(1)$ and $\Psi_\lambda(1)$ are known to have only simple zeros,* whilst $\Phi_\lambda(1)$ and $\Phi'_\lambda(1)$ do not both vanish for the same value of λ, nor do $\Psi_\lambda(1)$ and $\Psi'_\lambda(1)$.

* Since there is no non-zero solution of (3) for which $V_\mu(0)$ and $V'_\mu(0)$ vanish simultaneously, it is evident that the condition $V_\mu(0) = 1$ can be satisfied provided that the remaining conditions of (3) to (5) are compatible. For the required properties of μ, see Bôcher, 'Leçons sur les Méthodes de Sturm,' Chap. III. The properties also follow from Integral Equation Theory; see, *e.g.*, Goursat, 'Cours d'Analyse,' Chap. XXXIII (Paris, 1915), or Hilbert, 'Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen,' Kap. VII (Leipzig, 1912).

$\phi_\lambda(x)$ and $\psi_\lambda(x)$, regarded as functions of λ for any x in our range, are meromorphic with poles at $\lambda = \nu_n - i\gamma$ and $\lambda = \nu_n + i\gamma$ respectively.

Let us put

$$N(\lambda) = \tfrac{1}{2}\int_0^1 [\phi_\lambda(x) + \psi_\lambda(x)]\,dx. \tag{8}$$

Thus

$$N(\lambda) = \tfrac{1}{2}\int_0^1 \left[\frac{-2\lambda}{\lambda^2+\gamma^2} - \frac{d}{dx}\left\{ \frac{\eta(x)\,\Phi'_\lambda(x)}{(\lambda+i\gamma)^2\,\Phi_\lambda(1)} + \frac{\eta(x)\,\Psi'_\lambda(x)}{(\lambda-i\gamma)^2\,\Psi'_\lambda(1)} \right\} \right] dx,$$

i.e.,

$$N(\lambda) = -\left\{ \frac{\lambda}{\lambda^2+\gamma^2} + \frac{\eta(1)}{2(\lambda+i\gamma)^2}\cdot\frac{\Phi'_\lambda(1)}{\Phi_\lambda(1)} + \frac{\eta(1)}{2(\lambda-i\gamma)^2}\cdot\frac{\Psi'_\lambda(1)}{\Psi_\lambda(1)} \right\}. \tag{9}$$

This is valid for all λ except the poles of $\phi_\lambda(x)$ and $\psi_\lambda(x)$. This expression, however, does not define $N(\lambda)$ when $\lambda = \pm i\gamma$. In this case we must return to equations (7A), when for $\lambda = i\gamma$

$$N(\lambda) = \tfrac{1}{2}\int_0^1 \left\{ \phi_\lambda(x) + \int_x^1 [t/\eta(t)]\,dt \right\} dx$$

$$N(\lambda) = \frac{-1}{2}\left\{ \frac{1}{\lambda+i\gamma} + \frac{\eta(1)}{(\lambda+i\gamma)^2}\cdot\frac{\Phi'_\lambda(1)}{\Phi_\lambda(1)} - \int_0^1 \frac{x^2}{\eta(x)}\,dx \right\}. \tag{9A}$$

(9) and (9A), together with a similar expression for $N(\lambda)$ when $\lambda = -i\gamma$, show that $N(\lambda)$ is meromorphic, having only simple poles, these being at $\lambda = \nu_n \pm i\gamma$ $(n = 1, 2, 3, \ldots)$.

The simple roots of the equation

$$N(\lambda) = \alpha\lambda, \tag{10}$$

where α is any prescribed real number, we shall call *characteristic constants*, and the corresponding functions $\phi_\lambda(x)$, $\psi_\lambda(x)$, *characteristic functions*.

It is in terms of such functions that $f(x)$ and $g(x)$ are to be expanded.

§ 4. *Some General Properties of the Characteristic Constants and* $N(\lambda)$.

We shall now show that :—

(*a*) The characteristic constants which are not real occur in conjugate pairs.

(*b*) Complex characteristic constants have positive real parts.

(*c*) When $\alpha > 0$ there is no real negative characteristic constant, but when $\alpha < 0$ there may be one or more.

We use λ, λ' for any two values of our parameter, and $\bar{\lambda}$ for the conjugate of λ.

Evidently, from equations (1) and (2), $\phi_\lambda(x)$ and $\psi_\lambda(x)$ regarded as functions of λ are conjugate; and so are $\phi_{\lambda'}(x)$ and $\psi_\lambda(x)$. Thus it follows from the definition of $N(\lambda)$ that $N(\lambda)$ and $N(\lambda)$ are conjugate. It follows from this, in virtue of the fact that α is real, that complex characteristic constants occur in conjugate pairs. We see that $N(\lambda)$ is real when λ is real.

$$(\lambda + i\gamma)\int_0^1 \phi_\lambda(x)\,\phi_{\lambda'}(x)\,dx = \int_0^1 \phi_{\lambda'}(x)\{-1 - d/dx\,[\eta(x)\,.\,d\phi_\lambda(x)/dx]\}\,dx,$$

thus

$$(\lambda + i\gamma)\int_0^1 \phi_\lambda(x)\,\phi_{\lambda'}(x)\,dx = \int_0^1 \eta(x)\,\phi_\lambda'(x)\,\phi'_{\lambda'}(x)\,dx - \int_0^1 \phi_{\lambda'}(x)\,dx,$$

also

$$(\lambda' + i\gamma)\int_0^1 \phi_\lambda(x)\,\phi_{\lambda'}(x)\,dx = \int_0^1 \eta(x)\,\phi'_\lambda(x)\,\phi'_{\lambda'}(x)\,dx - \int_0^1 \phi_\lambda(x)\,dx,$$

hence

$$(\lambda - \lambda')\int_0^1 \phi_\lambda(x)\,\phi_{\lambda'}(x)\,dx = \int_0^1 [\phi_\lambda(x) - \phi_{\lambda'}(x)]\,dx,$$

also

$$(\lambda - \lambda')\int_0^1 \psi_\lambda(x)\,\psi_{\lambda'}(x)\,dx = \int_0^1 [\psi_\lambda(x) - \psi_{\lambda'}(x)]\,dx.$$

Hence

$$N(\lambda) - N(\lambda') = \tfrac{1}{2}(\lambda - \lambda')\int_0^1 [\phi_\lambda(x)\,\phi_{\lambda'}(x) + \psi_\lambda(x)\,\psi_{\lambda'}(x)]\,dx. \qquad (11)$$

Divide by $\lambda - \lambda'$ and let $\lambda' \to \lambda$, thus

$$N'(\lambda) = \tfrac{1}{2}\int_0^1 [\{\phi_\lambda(x)\}^2 + \{\psi_\lambda(x)\}^2]\,dx. \qquad (12)$$

In particular, if λ_p and λ_q are any two distinct characteristic constants, and $\phi_p(x)$, $\phi_q(x)$, etc., the corresponding characteristic functions,

$$\tfrac{1}{2}\int_0^1 [\phi_p(x)\,\phi_q(x) + \psi_p(x)\,\psi_q(x)]\,dx = \alpha. \qquad (13)$$

$$\tfrac{1}{2}\int_0^1 [\{\phi_p(x)\}^2 + \{\psi_q(x)\}^2]\,dx = N'(\lambda_p). \qquad (14)$$

We have further

$$(\lambda + i\gamma)\int_0^1 \phi_\lambda(x)\,\psi_{\lambda'}(x)\,dx = \int_0^1 \eta(x)\,\phi'_\lambda(x)\,\psi'_{\lambda'}(x)\,dx - \int_0^1 \psi_{\lambda'}(x)\,dx$$

$$(\lambda' - i\gamma)\int_0^1 \phi_\lambda(x)\,\psi_{\lambda'}(x)\,dx = \int_0^1 \eta(x)\,\phi'_\lambda(x)\,\psi'_{\lambda'}(x)\,dx - \int_0^1 \phi_\lambda(x)\,dx.$$

Hence

$$(\lambda + \lambda')\int_0^1 \phi_\lambda(x)\,\psi_{\lambda'}(x)\,dx = 2\int_0^1 \eta(x)\,\phi'_\lambda(x)\,\psi'_{\lambda'}(x)\,dx - \int_0^1 [\phi_\lambda(x) + \psi_{\lambda'}(x)]\,dx.$$

and

$$(\lambda + \lambda')\int_0^1 \phi_{\lambda'}(x)\,\psi_\lambda(x)\,dx = 2\int_0^1 \eta(x)\,\phi'_{\lambda'}(x)\,\psi'_\lambda(x)\,dx - \int_0^1 [\phi_{\lambda'}(x) + \psi_\lambda(x)]\,dx.$$

Thus

$$(\lambda + \lambda') \int_0^1 [\phi_{\lambda'}(x)\,\psi_\lambda(x) + \phi_\lambda(x)\,\psi_{\lambda'}(x)]\,dx$$

$$= 2\int_0^1 \eta(x)\,[\phi'_{\lambda'}(x)\,\psi'_\lambda(x) + \phi'_\lambda(x)\,\psi'_{\lambda'}(x)]dx - 2\,[N(\lambda) + N(\lambda')]. \quad (15)$$

In particular

$$(\lambda_p + \lambda_q)^{-1} \int_0^1 \eta(x)\,[\phi'_p(x)\,\psi'_q(x) + \phi'_q(x)\,\psi'_p(x)]\,dx$$

$$= \alpha + \tfrac{1}{2}\int_0^1 [\phi_p(x)\,\psi_q(x) + \phi_q(x)\,\psi_p(x)]\,dx,$$

and by (13)

$$(\lambda_p + \lambda_q)^{-1} \int_0^1 \eta(x)\,[\phi'_p(x)\,\psi'_q(x) + \phi'_q(x)\,\psi'_p(x)]\,dx$$

$$= \tfrac{1}{2}\int_0^1 [\phi_p(x) + \psi_p(x)]\,[\phi_q(x) + \psi_q(x)]\,dx. \quad (16)$$

In equation (16) suppose λ_p and λ_q are conjugate. It follows that $\phi'_p(x)$ and $\psi'_q(x)$, $\phi'_q(x)$ and $\psi'_p(x)$, $[\phi_p(x) + \psi_p(x)]$ and $[\phi_q(x) + \psi_q(x)]$ are pairs of conjugate functions, and in particular their products are positive. Hence $(\lambda_p + \lambda_q)$ must be positive, not including zero, for $\phi_\lambda(x) + \psi_\lambda(x)$ does not vanish for all x in $0 \leqslant x \leqslant 1$, whatever λ may be. Thus complex characteristic constants have their real parts strictly positive.

In (15) suppose $\lambda' = \lambda$. Then $\frac{1}{2}[N(\lambda) + N(\lambda)] =$ real part of $N(\lambda)$. Moreover, the two integrals in (15) are essentially real and positive in this case. Thus when λ has its real part negative or zero, $N(\lambda)$ has its real part strictly positive. In particular $N(\lambda) > 0$ for real values of λ in $\lambda \leqslant 0$. It follows from this that (10) has no real negative solutions if $\alpha > 0$, but may have such solutions if $\alpha < 0$.

§ 5. *Auxiliary Functions for Large Values of* $|\lambda|$.

Before proceeding to examine the distribution of the characteristic constants in greater detail, it will be convenient to discuss the forms of $v_\mu(x)$ and $N(\lambda)$ for large values of $|\mu|$ and $|\lambda|$.

$$\frac{d^2}{dx^2}\left[\eta(x)\frac{dv_\mu(x)}{dx}\right] + \frac{\mu}{\eta(x)}\left[\eta(x)\frac{dv_\mu(x)}{dx}\right] = 0,$$

and

$$(d^2/dz^2)[v_\mu(x)] + \mu\eta(x)\,v_\mu(x) = 0,$$

where

$$z = \int_\beta^x [\eta(x)]^{-1}\,dx, \qquad \beta \text{ being a fixed real constant.}$$

In a paper in the ' Proc. Lond. Math. Soc.' (series 2, vol. 23, p. 428) H. Jeffreys finds the form of the approximate solutions of equations of this type for large values of $|\mu|$ liable to an error of the order $|\mu|^{-\frac{1}{2}}$ of the solutions. Having regard to our boundary conditions, these take the form

$$\left.\begin{aligned} \frac{dv_\mu(x)}{dx} &= \left[\frac{H}{\eta(x)}\right]^{\frac{3}{4}} \sin\left\{\int_0^x \left[\frac{\mu}{\eta(x)}\right]^{\frac{1}{2}} dx\right\} \\ v_\mu(x) &= \left[\frac{\eta(0)}{\eta(x)}\right]^{\frac{1}{4}} \cos\left\{\int_0^x \left[\frac{\mu}{\eta(x)}\right]^{\frac{1}{2}} dx\right\} \end{aligned}\right\}. \tag{17}$$

In the former equation, H is an arbitrary constant arising because we have found $dv_\mu(x)/dx$ from an equation obtained by differentiating that which $v_\mu(x)$ satisfies. H must be so determined as to give agreement with (17).

$$\frac{dv_\mu(x)}{dx} = -\frac{[\mu^2\eta(0)]^{\frac{1}{4}}}{[\eta(x)]^{\frac{3}{4}}} \sin\left\{\int_0^x \left[\frac{\mu}{\eta(x)}\right]^{\frac{1}{2}} dx\right\}. \tag{18}$$

Put $I = \int_0^1 [\eta(x)]^{-\frac{1}{2}}\, dx$. Then we have

$$-\frac{\eta(1)}{2\mu^2} \cdot \frac{v'_\mu(1)}{v_\mu(1)} = \frac{[\eta(1)]^{\frac{1}{2}}}{2\mu^{\frac{3}{2}}} \tan[I\mu^{\frac{1}{2}}]. \tag{19}$$

The poles of this function are given by $I\mu^{\frac{1}{2}} = (r + \frac{1}{2})\pi$, where r may take all sufficiently large integral values, which may be supposed positive. Thus for all sufficiently large r,

$$\nu_r = [(r + \tfrac{1}{2})/I]^2\pi^2,$$

and the distance between two consecutive poles is, to our order of approximation,

$$\nu_r - \nu_{r-1} = 2r\pi^2/I^2 = 2\pi\nu_r^{\frac{1}{2}}/I.$$

Note also that

$$\frac{d}{d\mu}\left[-\frac{\eta(1)}{2\mu^2} \cdot \frac{v'_\mu(1)}{v_\mu(1)}\right] = \frac{[\eta(1)]^{\frac{3}{2}}}{2\mu^{\frac{3}{2}}}\left[\frac{I}{2\mu^{\frac{1}{2}}} \sec^2(I\mu^{\frac{1}{2}}) - \frac{3}{2\mu} \tan(I\mu^{\frac{1}{2}})\right]. \tag{20}$$

§ 6. $N(\lambda)$ *for Large Real Values of* λ.

The value of the function (19) for $\mu = \nu_r + \kappa$, where $\kappa = \sigma + i\tau$, τ being small compared with $|\mu|^{\frac{1}{2}}$ but not necessarily small compared with γ, and σ being numerically less than $2\pi\nu_r^{\frac{1}{2}}I^{-1}$, is

$$\begin{aligned} \tfrac{1}{2}\mu^{-\frac{3}{2}}[\eta(1)]^{\frac{1}{2}} \tan(I\nu_r^{\frac{1}{2}} + \tfrac{1}{2}I\kappa\nu_r^{-\frac{1}{2}}) &= -\tfrac{1}{2}\mu^{-\frac{3}{2}}[\eta(1)]^{\frac{1}{2}} \cot(\tfrac{1}{2}I\kappa\nu_r^{-\frac{1}{2}}) \\ &= -\frac{[\eta(1)]^{\frac{1}{2}}}{2\mu^{\frac{3}{2}}} \cdot \frac{\cos\sigma_1\cosh\tau_1 - i\sin\sigma_1\sinh\tau_1}{\sin\sigma_1\cosh\tau_1 + i\cos\sigma_1\sinh\tau_1} \\ &= -\frac{[\eta(1)]^{\frac{1}{2}}}{2\mu^{\frac{3}{2}}} \cdot \frac{\sin\sigma_1\cos\sigma_1 - i\sinh\tau_1\cosh\tau_1}{\sin^2\sigma_1 + \sinh^2\tau_1}, \end{aligned} \tag{21}$$

where $\sigma_1 = \frac{1}{2} I \sigma \nu_r^{-\frac{1}{3}}$ is numerically less than π and $\tau_1 = \frac{1}{2} I \tau \nu_r^{-\frac{1}{3}}$ is small. We easily find the extreme values of the real part of this to be $\pm [\eta(1)]^{\frac{1}{2}}/[2I\tau\nu_r]$, corresponding to $\sigma = \pm\tau$, when σ varies with τ constant.

If in (21) we put in turn $\tau = \gamma$, $\tau = -\gamma$, we obtain the last two terms of $N(\lambda)$ for large, positive, real λ; that is, neglecting small terms,

$$N(\lambda) = -\frac{1}{\lambda} - \frac{[\eta(1)]^{\frac{1}{2}}}{2\lambda^{\frac{3}{2}}} \cdot \frac{2 \sin \sigma_1 \cos \sigma_1}{\sin^2 \sigma_1 + \sinh^2 (\gamma I/2\lambda)^{\frac{1}{2}}}.$$

This is always of order $1/\lambda$ for sufficiently large λ.

Hence we have the result that when $\alpha \neq 0$, there can only be a finite number of real, positive, characteristic constants. If $\alpha = 0$ the same result holds provided $[\eta(1)]^{\frac{1}{2}} . [I\gamma]^{-1} < 1$, whereas there would be an infinite number of such constants if $[\eta(1)]^{\frac{1}{2}} . [I\gamma]^{-1} > 1$. If, however, $\alpha = 0$ and $[\eta(1)]^{\frac{1}{2}} . [I\gamma]^{-1} = 1$, $N(\lambda)$ would have to be examined to the next approximation.

It is necessary, however, in the present problem, to have $\alpha \neq 0$, as we shall see in the sequel (see § 7).

Returning to equation (19) for the case when μ has a negative real part and a finite imaginary part, we find the value $\frac{1}{2}[\eta(1)]^{\frac{1}{2}} |\mu|^{-\frac{1}{2}}$ for the modulus of the right-hand side. Hence $N(\lambda) = -\lambda^{-1}$ to the first approximation for large $|\lambda|$. From this and the fact that $N(\lambda)$ is strictly positive for real λ less than or equal to zero, it follows that when $\alpha < 0$ and only then, there is at least one (therefore an odd number) real, negative, characteristic constant.

From (21) we see that in the range $\nu_r - 2\pi\nu_r^{\frac{1}{3}} I^{-1} \leqslant \lambda \leqslant \nu_r + 2\pi\nu_r^{\frac{1}{3}} I^{-1}$

$$N(\lambda) = -\frac{1}{\lambda} - \frac{[\eta(1)]^{\frac{1}{2}}}{\lambda^{\frac{3}{2}}} \cdot \frac{\sin \sigma_1 \cos \sigma_1}{\sin^2 \sigma_1 + \sinh^2 \tau'_1}$$

λ being real, and $\tau'_1 = \frac{1}{2} I \gamma \nu_r^{-\frac{1}{3}}$. In the range $(\nu_r - \gamma) < \lambda < (\nu_r + \gamma)$ the variation of $N(\lambda)$ to our order of approximation arises from the second term of the above expression.

Again, at

$$\lambda = \nu_r + \gamma, \qquad N(\lambda) < -\nu_r^{-1},$$

and at

$$\lambda = \nu_r - \gamma, \qquad N(\lambda) > -\nu_r^{-1}.$$

Hence the graph of $N(\lambda)$ for real values of λ has the following features :—

$N(\lambda)$ is always strictly positive for $\lambda \leqslant 0$ and tends to zero as $\lambda \to -\infty$.

$N(\lambda) \to 0$ as $\lambda \to +\infty$.

$N'(\lambda)$ is negative for some range in $(\nu_r - \gamma) < \lambda < (\nu_r + \gamma)$ and positive in some range in $(\nu_{r-1} + \gamma) < \lambda < (\nu_r - \gamma)$, ($\nu_r$ large).

$N(\lambda)$ for Real Values of λ.

The graphs illustrate the behaviour of $N(\lambda)$ for real λ (dotted curves give $-\lambda^{-1}$).

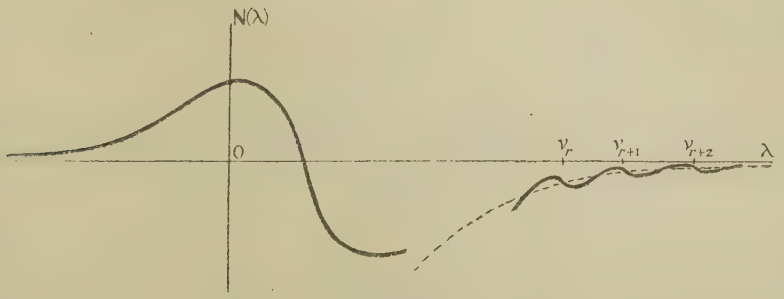


FIG. 1.—Case when $[\eta(1)]^{\frac{1}{2}} < I\gamma$.

$N(\lambda)$ has an odd finite number of real positive roots.

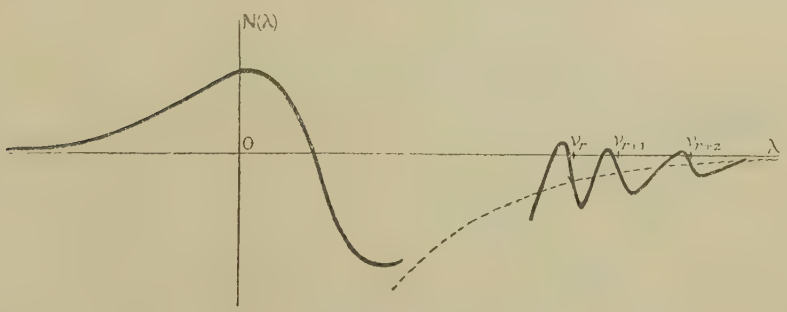


FIG. 2.—Case when $[\eta(1)]^{\frac{1}{2}} > I\gamma$.

$N(\lambda)$ has an infinity of real positive roots.

§ 7. *Characteristic Constants with* $|\lambda_p|$ *Large.*

We shall now show that for sufficiently large values of ν_r there is one characteristic constant within the circles $|\lambda - \nu_r \mp i\gamma| = \nu_r^{-\frac{1}{2}}$ provided $\alpha \neq 0$; and, further, that all the characteristic constants λ_p, for which $|\lambda_p|$ is sufficiently large, are so related to the poles of $N(\lambda)$ when $\alpha \neq 0$.

Putting $\kappa = \nu_r^{-\frac{1}{2}} e^{i\theta}$ and $\kappa = \pm 2i\gamma + \nu_r^{-\frac{1}{2}} e^{i\theta}$ in turn in (19) and (20), where κ is defined as in § 6, we find

$$N(\lambda) = -[\eta(1)]^{\frac{1}{2}} I^{-1} \nu_r^{-\frac{1}{2}} e^{-i\theta} [1 + 0(\nu_r^{-\frac{1}{2}})].$$

Differentiating (19) with respect to μ and making these substitutions we find

$$N'(\lambda) = [\eta(1)]^{\frac{1}{2}} I^{-1} e^{-2i\theta} [1 + 0(\nu_r^{-\frac{1}{2}})].$$

Now consider

$$\int \frac{N'(\lambda) - \alpha}{N(\lambda) - \alpha\lambda} \, d\lambda \quad \textit{taken round one of our circles.} \tag{22}$$

When $\alpha \neq 0$ this is

$$-i\int_0^{2\pi}\{\nu_r^{-\frac{1}{3}}I^{-1}\alpha^{-1}[\eta(1)]^{\frac{1}{2}}e^{-2i\theta}-\nu_r^{-\frac{2}{3}}\}e^{i\theta}[1+0(\nu_r^{-\frac{1}{2}})]\,d\theta \text{ having order } \nu_r^{-\frac{1}{2}}.$$

When $\alpha = 0$ this is

$$-i\int_0^{2\pi}[1+0(\nu_r^{-\frac{1}{2}})]\,d\theta = -2\pi i + 0(\nu_r^{-\frac{1}{2}}).$$

We know that for sufficiently large ν_r, each of these circles contains a simple pole of $N(\lambda) - \alpha\lambda$. Thus the above shows that they will contain also a simple zero of $N(\lambda) - \alpha\lambda$ if $\alpha \neq 0$ but no zero if $\alpha = 0$. (This is why we require $\alpha \neq 0$ to establish the expansions.)

Now consider the integral (22) taken round the circle $|\lambda| = r^2\pi^2 I^{-2}$, where r is a large integer. On this circle

$$\tan(I\lambda^{\frac{1}{2}}) = \tan(r\pi e^{\frac{1}{2}i\theta})$$

$$= \frac{\tan(r\pi\cos\frac{1}{2}\theta) + i\tanh(r\pi\sin\frac{1}{2}\theta)}{1 - i\tan(r\pi\cos\frac{1}{2}\theta)\tanh(r\pi\sin\frac{1}{2}\theta)}.$$

This is finite for all values of θ except possibly near such values as make

$$r\pi\cos\tfrac{1}{2}\theta = \pm(s+\tfrac{1}{2})\pi$$

where s is an integer $< r$ and $\geqslant 0$.
Here

$$r\pi\sin\tfrac{1}{2}\theta = \pm r\pi[1-(s+\tfrac{1}{2})^2/r^2]^{\frac{1}{2}}.$$

The least possible value of this is $\pi r^{\frac{1}{2}}$ for $s = r-1$ and $\tanh\pi r^{\frac{1}{2}} = 1$. Hence $\tan(r\pi e^{\frac{1}{2}i\theta})$ is finite for all values of θ, its upper bound not depending on r.

By formulæ (9) (19) (20) we have approximately

$$N'(\lambda) - \alpha = -\alpha$$

$$N(\lambda) - \alpha\lambda = -\alpha\lambda \quad \text{for sufficiently large } r.$$

Hence

$$\int_{|\lambda| = r^2\pi^2 I^{-2}} \frac{N'(\lambda)-\alpha}{N(\lambda)-\alpha\lambda}\,d\lambda = \int_0^{2\pi} i\,d\theta = 2\pi i.$$

Thus $N(\lambda) - \alpha\lambda$ has one more zero than poles in any sufficiently large circle about the origin. This establishes the second part of the theorem of this paragraph.

§ 8. *Change in Distribution of Characteristic Constants as* λ *Varies.*

Regarding α as a function of λ, we require to know the way in which α varies along the curves of real values of α.

$$\alpha = \lambda^{-1}N(\lambda). \tag{23}$$

Putting $\mu = M + iK$ in (19), where K is large and positive, and $MK^{-\frac{1}{2}}$ is finite or small, we get the value $2^{-\frac{3}{2}}(1-i)[\eta(1)]^{\frac{1}{2}}K^{-\frac{3}{2}}$

$$K^3\alpha = K + 2Mi - [\tfrac{1}{2}K\eta(1)]^{\frac{1}{2}}(1+i)[1+0(K^{-\frac{1}{2}})],$$

allowing for the possible error in (19).

Thus the curves of real α have no asymptotes parallel to the imaginary axis, but there is one branch extending to infinity in this direction, with a parabolic asymptote

$$8M^2 = \eta(1)K.$$

The real axis is a curve required, α passing through all negative values from $-\infty$ to -0 as λ decreases from -0 to $-\infty$. When $\lambda > 0$, α has the same sign as $N(\lambda)$, and since the variation of $N(\lambda)$ between successive maxima and minima is comparable with the value of the function at either of such points, while the variation of λ in such an interval is small compared with λ when λ is large, it follows that the maxima and minima of α occur for very nearly the same values of λ as those of $N(\lambda)$. Thus for large λ the general features of the variation of α (maxima, minima, zeros, sign) will be represented by the curves of figs. 1 and 2, the dotted curve giving $-\lambda^{-2}$ instead of $-\lambda^{-1}$.

By theorem (*b*), § 4, the curves (Γ say) of real α do not enter the complex part of the left-hand half plane. Also from equations (3) and (4) with μ real, the residues of $N(\lambda)$ at the poles are all real. Hence the tangent to the curve Γ at a pole of $N(\lambda)$ is parallel to the image in the real axis of the line joining the pole to the origin. Since the residue, when ν_r is large, is negative, $\alpha \to +\infty$ on the left and $\to -\infty$ on the right of such poles.

When a curve Γ crosses the real axis, $d\alpha/d\lambda = 0$ for α to be real along two distinct lines. We then have $\delta\alpha = \frac{1}{2}(\delta\lambda)^2 d^2\alpha/d\lambda^2$. The second derivative of α being real for real λ, $\delta\lambda$ must be real or purely imaginary according to the sign of $2\delta\alpha/(d^2\alpha/d\lambda^2)$. Such a curve crosses the real axis, perpendicularly, at each maximum and minimum of α. On leaving the real axis, describing a curve Γ,

α decreases if the point was a min. of α for real λ

α increases if the point was a max. of α for real λ.

In (21) put $\sigma_1 = \frac{1}{2}\pi$. We deduce that at $\lambda = \frac{1}{2}(\nu_r + \nu_{r+1}) + iy$ $N(\lambda) = \frac{1}{2}iI\lambda^{-2}[\eta(1)]^{\frac{1}{2}}y - \lambda^{-1}$, where $|\lambda|$ is large and y finite. For such values of λ, excluding λ real, α is not real.

The condition $d\alpha/d\lambda = 0$ for points of intersection of curves Γ with the real axis determines abscissæ of points at which the tangent to the curve $y = N(\lambda)$, (λ real) passes through the origin.

The curves of α real evidently consist of a series of closed curves and a single infinite branch approximating to $y = \pm 8[\eta(1)]^{-1}x^2$, where $\lambda = x + iy$, along which $\alpha \to +0$. Noting the behaviour of α near the loops for which λ is large, it is evident that this infinite branch crosses the real axis at a finite value of λ not zero. Whether there are closed branches to the left of it or not, it crosses at a point where α has a minimum value for real λ. If this value is negative, the branch must pass through a pole of $N(\lambda)$.

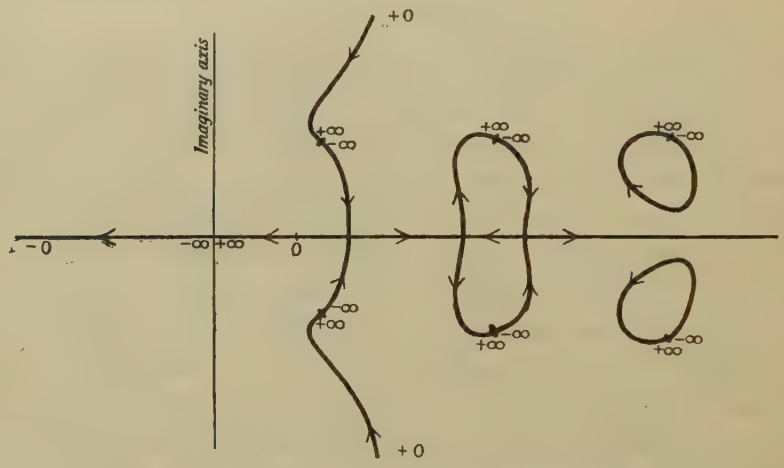


Fig. 3.—Graph illustrating Possible Form of Curves.

The arrows denote the directions in which α increases.

§ 9. *The Expansions.*

We have shown that, when $\alpha \neq 0$, there is a finite number of real characteristic constants (odd number) and an infinity of pairs of conjugate complex constants, and that for sufficiently large λ the latter are within $|\lambda|^{-\frac{1}{2}}$ of the poles of $N(\lambda)$, one to each. Now denote these by $\lambda_n (n = 1, 2, 3, \ldots)$ being ordered according to increasing $|\lambda|$ and then increasing arg λ. When convenient we shall denote conjugate characteristic constants by λ_s, $\lambda_{\bar{s}}$, s taking values of n for which λ_s has a strictly negative imaginary part. The

corresponding characteristic functions we denote by $\phi_n(x)$, $\phi_s(x)$, $\phi_{\bar{s}}(x)$, etc., respectively.

Suppose we can find sequences of numbers A_n, B_n $(n = 1, 2, \ldots)$ independent of x, such that

$$\left.\begin{aligned} f(x) &= \sum_{n=1}^{\infty} A_n \phi_n(x) & g(x) &= \sum_{n=1}^{\infty} A_n \psi_n(x) & 0 &= \sum_{n=1}^{\infty} A_n \\ 0 &= \sum_{n=1}^{\infty} B_n \phi_n(x) & 0 &= \sum_{n=1}^{\infty} B_n \psi_n(x) & 1 &= \sum_{n=1}^{\infty} B_n \end{aligned}\right\}. \tag{24}$$

Using equations (13) and (14) we find

$$\tfrac{1}{2}\int_0^1 [f(x)\,\phi_n(x) + g(x)\,\psi_n(x)]\,dx = [N'(\lambda_n) - \alpha]\,A_n$$

$$0 = \alpha + [N'(\lambda_n) - \alpha]\,B_n,$$

the series being supposed uniformly convergent, in $0 \leqslant x \leqslant 1$. Since $N'(\lambda_n) - \alpha \neq 0$ we may write

$$\left.\begin{aligned} A_n &= \frac{\tfrac{1}{2}\int_0^1 [f(x)\,\phi_n(x) + g(x)\,\psi_n(x)]\,dx}{N'(\lambda_n) - \alpha} \\ B_n &= \frac{-\alpha}{N'(\lambda_n) - \alpha} \end{aligned}\right\}. \tag{25}$$

We shall now substitute these values in the series (24) and examine the resulting series for convergence and sum.

§ 10. *Convergence.*

For sufficiently large values of s, the constant λ_s is equal to $\nu_{s'} + \kappa - i\gamma$, where $s' = \frac{1}{2}s$ (an integer) and

$$-\tfrac{1}{2}\nu_{s'}^{-\frac{1}{2}}[\eta(1)]^{\frac{1}{2}} \cdot \cot[\tfrac{1}{2}I\kappa\nu_{s'}^{-\frac{1}{2}}] = \alpha\lambda_s.$$

This gives

$$-\kappa = [\eta(1)]^{\frac{1}{2}}\alpha^{-1}I^{-1}(1/\nu_{s'})^2.$$

Whence also

$$N'(\lambda_s) = I\alpha^2[\eta(1)]^{-\frac{1}{2}}(\nu_{s'}^3) = N'(\lambda_{\bar{s}}),$$

and

$$\nu_{s'} = \pi^2(s+1)^2/4I^2.$$

From (7) and (17)

$$|\psi_{\bar{s}}(x)| = |\phi_s(x)| \leqslant 2|\alpha| \cdot [\eta(\xi)\,\eta(1)]^{-\frac{1}{4}}(\nu_{s'})^{\frac{3}{2}},$$

$$|\psi_s(x)| = |\phi_{\bar{s}}(x)| \leqslant I^{-1}\gamma^{-1}[\eta(\xi)/\eta(1)]^{-\frac{1}{4}}(\nu_{s'})^{-\frac{1}{2}},$$

where $\eta(\xi)$ is the minimum value of $\eta(x)$ in $(0, 1)$.

Hence for sufficiently large $n = s$ or $\bar{s}$

$$|A_n| < [\eta(1)/\eta(\xi)]^{\frac{1}{2}}I^{-1}|\alpha|^{-1}\left\{\left|\int_0^1 f(x)\,P(x)\,dx\right| + \left|\int_0^1 g(x)\,Q(x)\,dx\right|\right\}(\nu_{s'})^{-\frac{3}{2}},$$

where $P(x)$ and $Q(x)$ are bounded by ± 1. $f(x)$ and $g(x)$ being summable in (0, 1), the last bracket has a bound independent of $\nu_{s'}$.

$$|B_n| < [\eta(1)]^{\frac{1}{2}} I^{-1} |\alpha|^{-1} (\nu_{s'})^{-3},$$

$|B_n \phi_n(x)|$ and $|B_n \psi_n(x)|$ are both $< 2I^{-1} [\eta(1)/\eta(\xi)]^{\frac{1}{2}} (1/\nu_{s'}^{\frac{3}{2}})$.

Hence the series $\sum_{n=1}^{\infty} A_n$, $\sum_{n=1}^{\infty} B_n$, $\sum_{n=1}^{\infty} B_n \phi_n(x)$, $\sum_{n=1}^{\infty} B_n \psi_n(x)$ are absolutely and uniformly convergent, like Σn^{-3}, Σn^{-6}, Σn^{-3}, Σn^{-3}, respectively.

Now consider $A_n \phi_n(x)$ for large values of n. The alternate terms $A_s \phi_s(x)$ form a series converging absolutely, uniformly, and like Σn^{-4}. Hence our series converges or diverges with $\Sigma A_s \phi_s(x)$. For s large enough,

$$A_s \phi_s(x) = \frac{\frac{1}{2}\phi_s(x) \int_0^1 f(x)\, \phi_s(x)\, dx}{N'(\lambda_s) - \alpha} [1 + O(s^{-4})]$$

$$= 2I^{-1} [\eta(0)]^{-\frac{1}{2}} V_{s'}(x) \int_0^1 f(x)\, V_{s'}(x)\, dx\, [1 + O(s^{-1})],$$

allowing for the possible error to which our approximate solutions are subject. It is easy to show that

$$(\mu - \mu_1) \int_0^1 v_\mu(x)\, v_{\mu_1}(x)\, dx = \eta(1) [v_\mu(1)\, v'_{\mu_1}(1) - v_{\mu_1}(1)\, v'_\mu(1)].$$

Put $\mu_1 = \nu_r$. Then in virtue of equation (5)

$$(\mu - \nu_r) \int_0^1 v_\mu(x)\, v_{\nu_r}(x)\, dx = \eta(1) \cdot V'_r(1) [v_\mu(1) - v_{\nu_r}(1)].$$

Now let $\mu \to \nu_r$ and we find for all ν_r

$$\int_0^1 \{V_r(x)\}^2 dx = \eta(1)\, V'_r(1) [dv_\mu(1)/d\mu]_{\mu=\nu_r}. \tag{A}$$

For large values of r this takes the value $\frac{1}{2} I [\eta(0)]^{\frac{1}{2}}$.
Hence

$$A_s \phi_s(x) = \overline{V}_{s'}(x) \int_0^1 f(x)\, \overline{V}_{s'}(x)\, dx\, [1 + O(1/s')], \tag{26}$$

where $\overline{V}_{s'}(x)$ satisfies equations (3) to (5) for $\mu = \nu_{s'}$, except that $V_\mu(0) = 1$ is replaced by

$$\int_0^1 \left\{\overline{V}_{s'}(x)\right\}^2 dx = 1. \tag{27}$$

Now $\Sigma \overline{V}_r(x) \int_0^1 f(x)\, \overline{V}_r(x)\, dx$ is the Sturm-Liouville expansion of $f(x)$ corresponding to the above conditions, and this we have assumed to converge to $f(x)$ in (0, 1).

Moreover, the term $O(1/s')$ in (26) could, by the nature of our approximations, be put in the form $O(1/s') = a/s + b/s^2 + {}_0(s^{-2})$ where, a and b are fixed. Whence it is easily deduced in the case when $a \neq 0$ that this factor decreases steadily to zero. In any case it follows that

$$\Sigma \bar{V}_s(x) \int_0^1 f(x)\, \bar{V}_s(x)\, .\, dx\, O(s^{-1}) \quad \text{converges.}$$

Whence

$$\Sigma A_s \phi_s(x) \quad \text{converges.}$$

Therefore

$$\Sigma A_n \phi_n(x) \quad \text{converges.}$$

And similarly $\Sigma A_n \psi_n(x)$ converges if $g(x)$ has a Sturm-Liouville expansion of the same type as $f(x)$.

§ 11. *Values of Series.*

Consider the contour integrals :—

$$I_1 = \int_C \frac{-\alpha}{N(\lambda) - \alpha\lambda}\, d\lambda, \qquad I_2 = \int_C \left\{ \frac{\frac{1}{2}\int_0^1 [f(x)\, \phi_\lambda(x) + g(x)\, \psi_\lambda(x)]\, dx}{N(\lambda) - \alpha\lambda} \right\} d\lambda,$$

$$J_1 = \int_C \frac{-\alpha\phi_\lambda(x)}{N(\lambda) - \alpha\lambda}\, d\lambda, \qquad J_2 = \int_C \left\{ \frac{\frac{1}{2}\phi_\lambda(x) \int_0^1 [f(x)\, \phi_\lambda(x) + g(x)\, \psi_\lambda(x)]\, dx}{N(\lambda) - \alpha\lambda} \right\} d\lambda,$$

$$K_1 = \int_C \frac{-\alpha\psi_\lambda(x)}{N(\lambda) - \alpha\lambda}\, d\lambda, \qquad K_2 = \int_C \left\{ \frac{\frac{1}{2}\psi_\lambda(x) \int_0^1 [f(x)\, \phi_\lambda(x) + g(x)\, \psi_\lambda(x)]\, dx}{N(\lambda) - \alpha\lambda} \right\} d\lambda,$$

where C is the circle $|\lambda| = r^2\pi^2 I^{-2}$ for large r.

We have seen in § 7 that $N(\lambda)$ can be neglected compared with $\alpha\lambda$ on this circle, on which there are no poles of any of the above integrands ; and also $\phi_\lambda(x)$ and $\psi_\lambda(x)$ are of the order λ^{-1} on it by (7) and (17). Hence all these integrals except I_1 tend to zero as $r \to \infty$ and $\lim_{r \to \infty} I_2, = 2\pi i$.

All these integrands have simple poles at the simple zeros of $N(\lambda) - \alpha\lambda$, *i.e.*, for all values of λ which are characteristic constants, and the corresponding residues are obtained by writing the appropriate values of λ (say λ_n) in the numerators, and replacing the denominators by $N'(\lambda_n) - \alpha$. The only possibility of other simple poles arises through the poles of $\phi_\lambda(x)$ and $\psi_\lambda(x)$. By (7) and (9) we see that at an infinity of $\phi_\lambda(x)$ or of $\psi_\lambda(x)$, $N(\lambda)$ is infinite to the same order ; hence no such poles arise in the integrands of J_1, K_1, or I_2, but they will arise in the case of J_2 at the poles of $\phi_\lambda(x)$, and in the case of K_2 at the poles of $\psi_\lambda(x)$.

Consider J_2. Near any pole arising in this manner, say $\lambda = \nu_n - i\gamma$, we have for the form of the integrand near $\mu = \nu_n$,

$$\frac{v_\mu(x)\int_0^1 f(x)\,v_\mu(x)\,dx}{2\,[\mu v_\mu(1)]^2\,N(\mu - i\gamma)} = \frac{v_\mu(x)\int_0^1 f(x)\,v_\mu(x)\,dx}{-\eta\,(1)\,v_\mu(1)\,v'_\mu(1)} = -\frac{V_n(x)\int_0^1 f(x)\,V_n(x)\,dx}{\delta\mu\,.\,\eta\,(1)\,V'_n(1)\,[dv_\mu(1)/d\mu]_{\mu=\nu_n}}.$$

Hence, using the result (A) of § 10, we find for the residue at this pole

$$-\frac{\overline{V}_n(x)\int_0^1 f(x)\,\overline{V}_n(x)\,dx}{\int_0^1 \{\overline{V}_n(x)\}^2\,dx},$$

which is the general term of a Sturm-Liouville expansion of $-f(x)$.

Hence evaluating these integrals by the theorem of residues also and letting $r \to \infty$, we find that the equations (24) are valid, the coefficients A_n, B_n, being given by the equations (25).

Hence if $f(x)$ and $g(x)$ have Sturm-Liouville expansions in (0, 1) in terms of the functions defined by the equations

$$\left.\begin{aligned} &\frac{d}{dx}\left[\eta(x)\,\frac{d\overline{V}_{s'}(x)}{dx}\right] + \nu_{s'}\overline{V}_{s'}(x) = 0 \\ &\overline{V}'_{s'}(0) = 0 \qquad\qquad \overline{V}_{s'}(1) = 0 \\ &\qquad\qquad \int_0^1 \{\overline{V}_s(x)\}^2 dx = 1 \end{aligned}\right\},$$

then they can be expanded in terms of $\phi_n(x)$ and $\psi_n(x)$ respectively, solutions of (1) when λ is a simple root of (10), in such form that the same sequence of coefficients occurs in each expansion, and this sequence forms a series with any sum K we please. The expansions have the form

$$f(x) = \sum_{n=1}^{\infty} C_n \phi_n(x), \qquad g(x) = \sum_{n=1}^{\infty} C_n \psi_n(x),$$

where $C_n = A_n + KB_n$, A_n and B_n being defined by the equations (25).

The Expansion of Charcoal on Sorption of Carbon Dioxide.

By F. T. Meehan, Ph.D., B.Sc. (of the Building Research Station).

(Communicated by Sir William Hardy, F.R.S.—Received February 23, 1927.)

In the examination of certain building materials it has been found that an expansion takes place on absorption of water. This expansion is a movement of the order of strain movements in structural materials and not of the magnitude expected by the physical chemist in connection with elastic gels. Accompanying this small expansion of building materials is a fundamental change in physical properties.

McBain ('Journal of Physical Chemistry') showed that the building materials which possess the property of swelling with increased water content are colloidal structures, presumably of the rigid gel type. It therefore appeared desirable to ascertain whether rigid gels in general exhibited this small swelling movement. Graham* has shown that palladium expands linearly by 1·6 per cent. when it absorbs hydrogen.

Sorption of gas by charcoal was chosen for experiment because much work had already been done on it in other directions.

It has been found that charcoal does expand on sorption of gas, and that this expansion is of the same order as the moisture expansion in building materials. This discovery may have a wider significance than that required immediately by the Building Research Station; it may be that the type of expansion is characteristic of rigid gels.

Preliminary determinations were made by a volumetric method. The volumes of sticks of birch and beech charcoals found by displacement of mercury before and after absorption of carbon dioxide at atmospheric pressure showed that a volume change of the order of 0·5 per cent. took place at 15° C. The method was unsatisfactory because mercury entered the specimen. It was thus not possible to evacuate the air and moisture; moreover, some mercury remained in the charcoal and covered the surface. It was accordingly decided to adopt the method of linear measurement described below.

* 'Roy. Soc. Proc.,' A, vol. 15, p. 223 (1867).

Experimental Method.

Preparation of Charcoal.

Three-inch cubes of yellow pine* were slowly dried and then placed in a charcoal packed cylinder having small gas vents in its ends. The cylinder was heated uniformly for approximately four hours in a muffle furnace to about 500°. After gradual cooling, the charcoal so obtained was razor-trimmed into blocks approximating to 2-inch cubes, with one pair of faces cut as nearly as possible parallel to the grain; another, normal to the grain; and the third, tangential to the annular rings.

Expansion Apparatus.

The apparatus used for measuring the expansion is shown in fig. 1. A cube of charcoal was placed in the extensometer E commonly in use at the Building

Fig. 1.

Research Station. It is of the optical lever type, fitted with a screw S for calibration purposes. The extensometer, a thermometer and a hygroscope (found necessary to indicate minimum humidity of the system since ingress of dry gas caused a contraction of imperfectly dried charcoal) were placed in the vacuum chamber A. This chamber was a steel cylinder, 12 inches high and 8 inches in diameter, standing on a rubber ring let into a brass base plate.

* Several other woods were tried but none gave pieces of charcoal of suitable dimensions. No failures were experienced with yellow pine.

The mirror of the extensometer was set parallel to an optically plane glass window B in the vacuum chamber and the instrument calibrated, the reflection in the mirror of an illuminated millimetre scale being observed through a telescope fitted with cross hairs. The scale was about 8 feet from the mirror and a movement of the specimen of 0·001 inches corresponded to about 75 mm. on the scale. The top of the chamber was next fixed in position; it was similar to the brass base, to which it was bolted in three places, but had three inlet tubes (i) the gas inlet which penetrated well into the chamber to ensure circulation, (ii) the connection to a "Hyvac" pump giving a vacuum of 0·01 mm. Hg, and (iii) the connection to a mercury manometer registering pressures up to 800 mm. By alternate evacuations and admissions of dry air the charcoal was dried until a steady scale reading was reached at the lowest pressure attainable. The carbon dioxide was obtained from a cylinder whence it was passed through sulphuric acid washing bottles and a phosphorus pentoxide tube into the chamber. Equilibrium was usually reached about two hours after each alteration of the pressure.

Experimental Results.

The temperature coefficient of the apparatus was found to be 0·000023 inch per degree Centigrade and in the earlier experiments the temperature was maintained constant to $\pm 0{\cdot}5°$ C.

The first measurements taken were the expansions in the three grain directions due to one atmosphere of carbon dioxide. Tables I and II show the results.

Table I.

Temperature $16{\cdot}4 \pm 0{\cdot}5°$ C.

Direction.	Length of edge.	Expansion.	Per cent. linear expansion per atmosphere CO_2.
	inches.	inches $\times 10^{-3}$	
Along grain	2·125	4·286	0·2017
Radially	2·140	4·890	0·2285
Tangentially	2·030	4·696	0·2314

Cubical expansion per atmosphere $CO_2 = 0{\cdot}6616$ per cent.

Table II.

Temperature 10·0 ± 0·5° C.

Direction.	Length of edge.	Expansion.	Per cent. linear expansion per atmosphere CO_2.
	inches.	inches × 10^{-3}	
Along grain	2·100	5·600	0·2667
Radially	2·110	4·800	0·2275
Tangentially	2·100	5·300	0·2524

Cubical expansion per atmosphere CO_2 = 0·7466 per cent.

The variation of expansion with pressure of gas was next measured at constant temperature and in these experiments the temperature was electrically controlled to be constant to ± 0·1° C. The same piece of charcoal was used at four different temperatures and all the measurements were made along the grain. To ensure that the low pressure readings were recorded when the charcoal contained the correct quantity of carbon dioxide at each temperature, the procedure adopted was to start with a gas pressure of over one atmosphere, take the equilibrium pressure and scale readings at the experimental temperature, decrease the pressure by about 100 mm. Hg and again read at equilibrium. The pressure was so brought in stages to the minimum at which the scale was read. Then the chamber was refilled to about atmospheric pressure and the next temperature adjustment made.

The results obtained are given in Tables III to VI and are also plotted together in fig. 2. From the curves, the expansions for pressures of 760 mm. Hg were estimated for purposes of comparison.

Table III.

Temperature 15·0 ± 0·1° C.

Pressure of CO_2 (mm. Hg).	Percentage linear expansion.
760	0·1899*
735	0·1865
652	0·1723
570	0·1572
459	0·1371
373	0·1193
295	0·1017
200	0·0777
103	0·0435
1	0·0000

Table IV.

Temperature 23·0 ± 0·1° C.

Pressure of CO_2 (mm. Hg).	Percentage linear expansion.
775	0·1492
760	0·1466*
661	0·1349
544	0·1175
429	0·0999
355	0·0875
222	0·0616
144	0·0431
1	0·0000

* Estimated.

Table V.

Temperature 27·9 ± 0·1° C.

Pressure of CO_2 (mm. Hg).	Percentage linear expansion.
760	0·1292*
757	0·1288
674	0·1196
589	0·1085
491	0·0962
400	0·0832
294	0·0666
197	0·0487
99	0·0271
1	0·0000

Table VI.

Temperaure 35·8 ± 0·1° C.

Pressure of CO_2 (mm. Hg).	Percentage linear expansion.
770	0·1122
760	0·1100*
669	0·0888
558	0·0777
465	0·0641
359	0·0493
260	0·0401
200	0·0228
103	0·0126
1	0·0000

* Estimated.

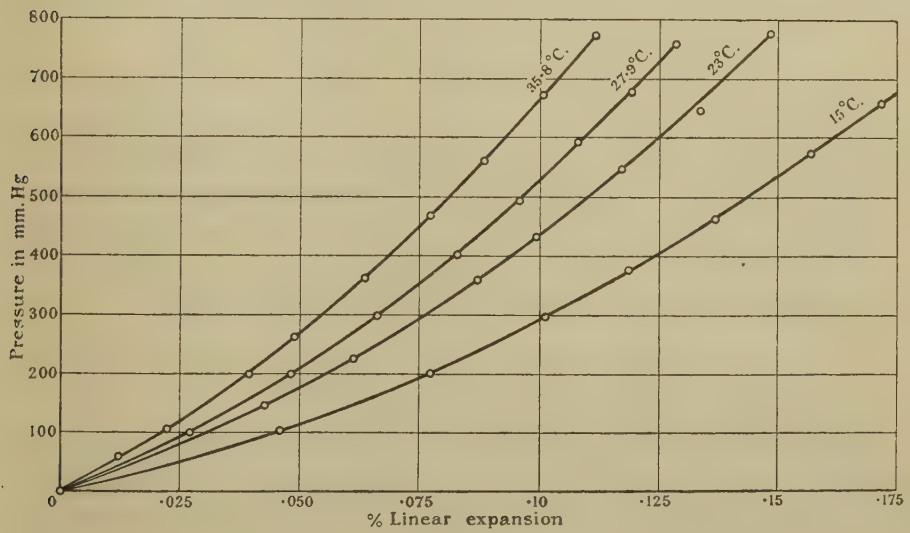


Fig 2.

In this series and in that described earlier it was found that the same equilibrium points were reached irrespective of whether the gas pressure had been arrived at from above or below, showing that, over the pressure range examined, the process was reversible.

Discussion of Results.

From the experiments first described it will be seen that the expansion is independent of the direction of the grain; hence it must be of a different nature from the expansion of wood due to the absorption of moisture as found

in Robertson's experiments.* This uniformity of expansion in different directions indicates that the æolotropy of wood is destroyed on carbonisation and that wood charcoal is isotropic. None of the experiments indicated that the charcoal was permanently deformed, and since the results are independent of the direction of the change of gas pressure, the variation in volume is reversible.

A quantitative examination of the variation of expansion with pressure is interesting. At constant temperature the relation must be

$$E_{obs.} = f(p) - Cp,$$

where $E_{obs.}$ is the observed expansion, C the compressibility of the charcoal and p the gas pressure. In general C† is of the order 10^{-6} per atmosphere and so is negligible in the present work, whence we have

$$E_{obs.} = f(p).$$

For the purpose of determining the form of the function, $f(p)$, two expressions have been examined. The first

$$E = kp^{1/n} \qquad (1)$$

based on the usual adsorption expression, where E is the percentage linear expansion, p the pressure of carbon dioxide in millimetres Hg and k and n constants, functions of temperature. From this expression the relation of log E to log p should be linear. That the expression does not hold accurately is shown by the curves in fig. 3. The second expression tested is based on the adsorption equation due to A. M. Williams‡ written

$$\log E/p = G + HE, \qquad (2)$$

where E and p are as before, G and H constants.

Equation (2) implies that log E/p when plotted against E would give a straight line, the necessary constants being functions of the temperature. That equation (2) more nearly represents the observations than equation (1) is shown by the curves in fig. 4.

Equation (1) may be written as

$$E/p = kp^{(1/n-1)}$$

from which log E/p is linearly related to log p. For purposes of comparison log

* "Report on the Materials of Construction used in Aircraft and Aircraft Engines," by C. F. Jenkin ('Aeronautical Research Committee, 1920').

† *Cf.* Richards, 'J. Chem. Soc.,' 1911, also Bridgman, 'Proc. Nat. Acad. Sci.,' 1915, *et seq.* C for graphite is 3×10^{-6} per bar.

‡ 'Roy. Soc. Proc., Edin.,' vol. 38, p. 23 (1918); vol. 39, p. 48 (1919); 'Roy. Soc. Proc.,' A, vol. 96, pp. 287, 298 (1919).

E/p has been plotted against E and against log p on the same figure and to commensurate scales.

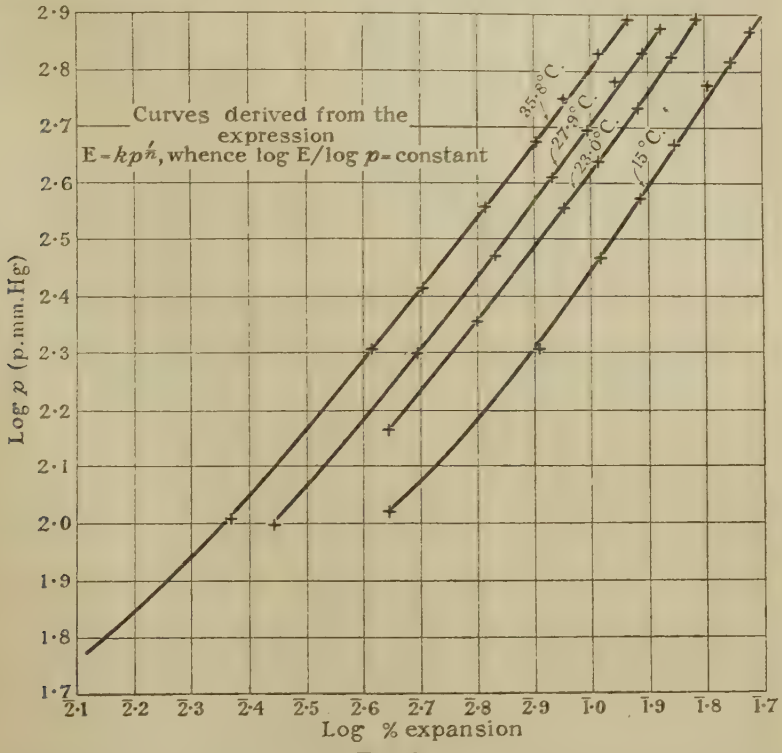


Fig. 3.

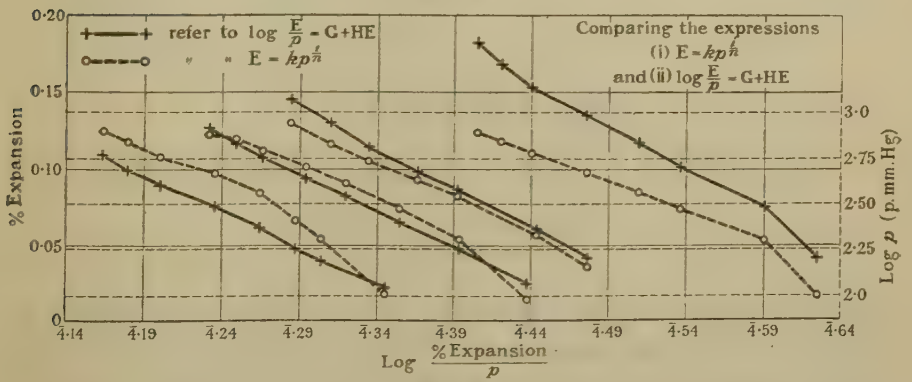


Fig. 4.

Assuming that the lines on fig. 4 derived from equation (2) are parallel, by calculation from the data equation (2) becomes

$$E = A' - \tfrac{1}{2} \log E/p$$

where A′ has the value $-0{\cdot}14 - 0{\cdot}175\ T\ (100 - T)10^{-3}$.

That is, where E is percentage linear expansion, p the gas pressure in millimetres Hg, and T the temperature on the Centigrade scale.

$$E = -0{\cdot}14 - 0{\cdot}175\,T\,(100 - T)10^{-3} - \tfrac{1}{2}\log_{10} E/p.$$

The relation of expansion to temperature at constant pressure is shown by the curves on fig. 5 to be

$$1/E = AT + B \qquad (3)$$

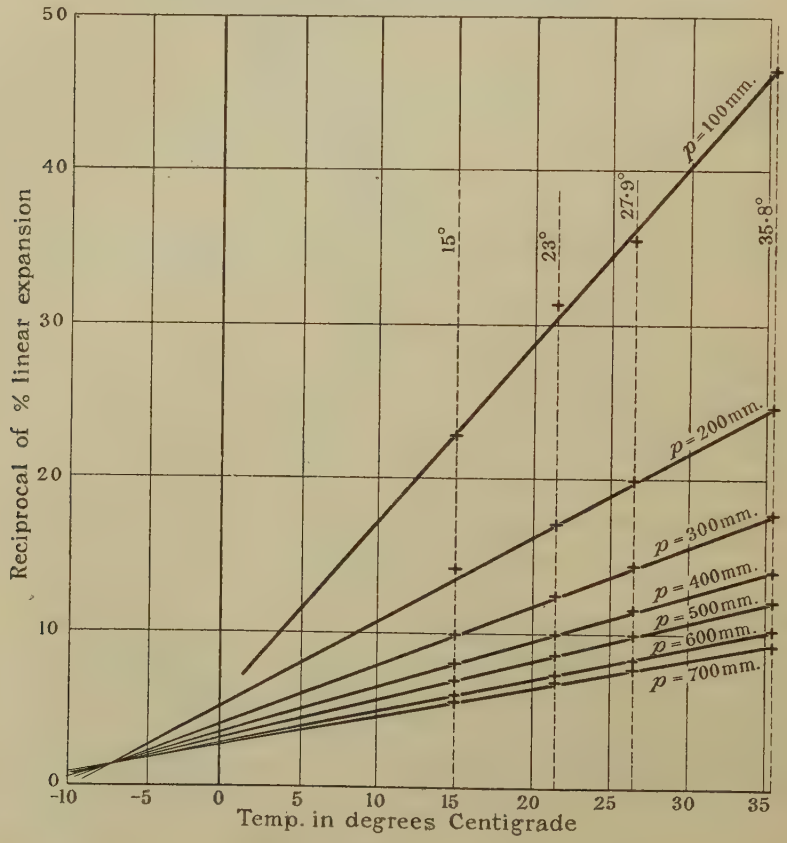


Fig. 5.

where E is percentage linear expansion, T the temperature Centigrade, A and B constants. That this relation holds over the whole pressure range examined may be seen from the linearity of the curves shown at pressure intervals of 100 mm. Hg. The constants A and B are both functions of pressure but the intersection of the lines shows that there is only one independent constant.

By calculation from the curves of fig. 5 equation (3) reduces to the form

$$(T + 6{\cdot}5)\,(100/p)\,0{\cdot}885 = (1/E - 1{\cdot}59),$$

where T is the temperature in degrees C, p the pressure in millimetres Hg, and E the percentage linear expansion.

It may be noted that the expression (2) and (3) hold continuously for temperatures above and below 31·5° C., the critical temperature of carbon dioxide.

Summary.

1. The expansion of yellow pine charcoal due to sorption of carbon dioxide up to a pressure of one atmosphere has been measured at various temperatures. The process is reversible.

2. The æolotropic structure of wood is destroyed on carbonisation, the resulting charcoal being isotropic.

3. At constant temperature the relation of expansion to pressure is better represented by the equation

$$\log E/p = G + HE$$

than by the usual expression $E = kp^{1/n}$.

4. At constant pressure expansion is related to temperature by the expression

$$1/E = AT + B.$$

5. The relations of expansion to temperature and pressure hold both above and below the critical temperature of carbon dioxide.

The author wishes to express his appreciation of the helpful suggestions and assistance given by his colleagues.

A Contribution to Modern Ideas on the Quantum Theory.

By H. T. Flint and J. W. Fisher, Wheatstone Laboratory, King's College, London.

(Communicated by O. W. Richardson, F.R.S.—Received February 21, 1927.)

I.—*The de Broglie Phase Wave in Generalised Space-Time.*

The relativity theory of gravitation indicates that space-time is a four dimensional continuum in which the line element is measured by the equation

$$(ds)^2 = g_{mn}dx_m dx_n, \tag{1}$$

the notation being that generally adopted.

The world-lines or natural tracks of free particles in this space are geodesics.

From (1) we have

$$g_{mn}\frac{dx_m}{ds}\cdot\frac{dx_n}{ds} = 1, \tag{2}$$

the quantity on the left being an expression corresponding to the kinetic energy of ordinary dynamics for a particle of unit mass. This correspondence is readily appreciated if it be noted that dx_m/ds is the natural extension of the velocity, dx_m/dt.

dx_m/ds ($m = 1, 2, 3, 4$) is one of the direction cosines of the track of the particle.

Equation (2) corresponds to

$$g_{mn}\frac{dx_m}{dt}\cdot\frac{dx_n}{dt} = \left(\frac{ds}{dt}\right)^2, \qquad m, n = 1, 2, 3,$$

in three dimensions.

Hence since ds/dt is the speed of the particle we could write

$$2\mathrm{T} = g_{mn}\frac{dx_m}{dt}\cdot\frac{dx_n}{dt}, \qquad m, n = 1, 2, 3,$$

where T is the kinetic energy.

Thus in a conservative system with total energy, E, and potential energy V, we have

$$g_{mn}\frac{dx_m}{dt}\cdot\frac{dx_n}{dt} = 2(\mathrm{E} - \mathrm{V}), \qquad m, n = 1, 2, 3. \tag{3}$$

This point is mentioned to show what is meant by the statement that the left side of (2) corresponds to the kinetic energy in ordinary mechanics and

it appears that instead of 2T or $\frac{2}{\cdot}(E - V)$ we have here to do with the invariant expression of (2).

We may define as the momentum of our unit mass the covariant vector whose typical component, $p_m = g_{m\alpha}\frac{dx_\alpha}{ds}$.

Then (2) is replaced by

$$g^{mn}p_m p_n = 1, \tag{4}$$

where again g^{mn} has the usual significance.

(4) is the expression of the invariant expression in terms of covariant quantities and corresponds to the expression of kinetic energy in terms of generalised momenta instead of velocities. We may make the substitution

$$p_m = \frac{\partial W}{\partial x_m},$$

and so express (4) as

$$g^{mn}\frac{\partial W}{\partial x_m}\frac{\partial W}{\partial x_n} = 1,$$

which is equivalent to $|\text{ grad. } W|^2 = 1$* where the gradient operator is to be understood as the generalised gradient of the analysis of tensors.

Now surfaces, W = constant, have normals whose direction cosines are $\frac{\partial W}{\partial x^m}$ $(m = 1, 2, 3, 4)$ and since from (2) we have

$$p_n\frac{dx_n}{ds} = 1,$$

it follows that

$$\frac{\partial W}{\partial x^n}\cdot\frac{dx_n}{ds} = 1,$$

i.e., the geodesic is normal to the surfaces.

It becomes possible to regard the continuum as filled with geodesics and a family of surfaces which bear to each other the same relation as lines of force and equipotential surfaces in an electrostatic field.

This mutual relation appears to be the basis of de Broglie's view that the motion of a particle can be associated with a phase-wave. The particle in the continuum may be located as a point of the geodesic or it may be associated with the particular W-surface through that point.

* *Cf.* "Quantisierung als Eigenwertproblem," Shroedinger, 'Ann. d. Physik,' vol. 79, p. 492 (1926).

To a three-dimensional observer the particle appears as a point travelling along its world line and to him the W-surface will appear as a wave.*

II.—*The Wave Equation as a Fundamental Law of Nature.*

The relativity theory of gravitation teaches that gravitational phenomena are the physical representation of the geometry of space-time described in terms of the components, g_{mn}, which fix the nature of the geometry. It would be a natural extension of these principles to proceed by a similar method to include the phenomena of electromagnetism. It would be very satisfactory if, by choosing appropriate components, g_{mn}, we could include in one system of geometry both gravitation and electromagnetism.

Natural phenomena do not appear to occur in accordance with such a simple scheme.

Weyl and Eddington have made the most natural extension to space-time geometry by the introduction of local scales of measurement, a standard for the measurement of $(ds)^2$ being appropriate to each point of the continuum. They have in this way found it possible to include the second class of phenomena in the scheme.

It is possible to make this inclusion in another way.

Let it be supposed that the energy, momentum and stresses associated with electromagnetic phenomena exert an influence upon the geometry of space, contributing along with gravitation to the components, g_{mn}.† Let the line-element be still measured by the rule,

$$(ds)^2 = g_{mn}\, dx_m\, dx_n.$$

This is not enough for the description of the phenomena. We must add a new element to the theory.

Let a four-vector, $\mathbf{s}'$, the current four-vector, be associated with elements of three-dimensional volume. The affix denotes the contravariant character of the vector and its components are (s^1, s^2, s^3, s^4). The three-dimensional volume element may be considered as a covariant vector with components $(d\sigma_1, d\sigma_2, d\sigma_3, d\sigma_4)$ and the integral

$$\iiint s^n\, d\sigma_n,$$

is of importance in the theory.

* [*Added March* 16, 1927.—We have found it possible to extend the view outlined here in a way which leads to the conception of a world-line possessing structure, from which one passes naturally to the idea of a quantum of action. We hope to communicate our conclusions shortly.]

† *Cf.* Eddington, 'The Mathematical Theory of Relativity,' p. 185.

This integral may be shown to be equivalent to the four-dimensional integral

$$\iiiint \text{div. } \mathbf{s}' \, . \, dv.$$

In the case when $\mathbf{s}'$ is the current four-vector div. $\mathbf{s}'$ vanishes. Associated with the current four-vector is a six-vector, which we may regard as a tensor of the second order, known as the Minkowski six-vector. If we denote its typical components by F_{mn} where

$$\mathrm{F}_{mm} = 0, \qquad \mathrm{F}_{mn} = \mathrm{F}_{nm}$$

we have

$$\iint \tfrac{1}{2} \mathrm{F}_{mn} \, d\mathrm{A}^{mn} = \iiint (\text{Lor. } \mathbf{F})^n \, d\sigma_n,$$

where $d\mathrm{A}^{mn}$ is a typical component of the contravariant element of area.

The Lorentz operator is to be understood in its general sense appropriate to generalised space-time.

There is a connection between $\mathbf{s}'$ and $\mathbf{F}$, for the equations

$$(\text{Lor. } \mathbf{F})^n = s^n \, (n = 1, 2, 3, 4)$$

give us half the Maxwell equations of electrodynamics in their most general form. The other half are given by

$$(\text{Lor. } \mathbf{G})^n = 0.$$

We might combine these two and write

$$\text{Lor. } (\mathbf{F} + \mathbf{G})^n = s^n,$$

if it were possible to dissociate $\mathbf{F}$ and $\mathbf{G}$ so as to reproduce the equations

$$(\text{Lor. } \mathbf{F})^n = s^n$$

$$(\text{Lor. } \mathbf{G})^n = 0.$$

The derivation of the general electrodynamical equations may be carried out by remembering that the components of the two reciprocal tensors are respectively

$$\mathrm{F}^{23} = \mathrm{H}_x, \ \mathrm{F}^{31} = \mathrm{H}_y, \ \mathrm{F}^{12} = \mathrm{H}_z, \ \mathrm{F}^{14} = -\mathrm{E}_x, \ \mathrm{F}^{24} = -\mathrm{E}_y, \ \mathrm{F}^{34} = -\mathrm{E}_z,$$

$$\mathrm{G}^{23} = \mathrm{E}_x, \ \mathrm{G}^{31} = \mathrm{E}_y, \ \mathrm{G}^{12} = \mathrm{E}_z, \ \mathrm{G}^{14} = \mathrm{H}_x, \ \mathrm{G}^{24} = \mathrm{H}_y, \ \mathrm{G}^{34} = \mathrm{H}_z.$$

We may note that the equations

$$\iint \tfrac{1}{2} \mathrm{G}_{mn} \, d\mathrm{A}^{mn} = 0,$$

and

$$\iiint s^n d\sigma_n = 0,$$

or what is the same thing

$$\text{Lor. } \mathbf{G} = 0$$

and

$$\text{div. } \mathbf{s}' = 0$$

are similar relations, applying to two and three-dimensions respectively. Lor. $\mathbf{G} = 0$ is not sufficient for the description of the electromagnetic phenomena, we require to add to it the equation

$$\text{Lor. } \mathbf{F} = \mathbf{s}'.$$

We shall suppose that in the same way div. $\mathbf{s}' = 0$ is also insufficient to cover all the phenomena possible in nature, and shall suppose that we must consider another four-vector $\mathbf{r}'$ whose divergence does not vanish but is equal to some quantity ψ which is, of course, a scalar.

We thus complete our equations by adding

$$\text{div. } \mathbf{r}' = \psi.$$

The system of equations so obtained is strikingly simple and uniform. It is

$$\left.\begin{aligned} \text{div. } \mathbf{s}' &= 0 \\ \text{div. } \mathbf{r}' &= \psi \end{aligned}\right\}$$

$$\left.\begin{aligned} \text{Lor. } \mathbf{G} &= 0 \\ \text{Lor. } \mathbf{F} &= \mathbf{s}' \end{aligned}\right\}$$

and to these we add the equation for the four-vector potential

$$\text{curl } \mathbf{a}_1 = \mathbf{F}.$$

In these equations, ψ may be regarded as the most fundamental quantity. It is to be associated with four-dimensional volume elements in the same way that $\mathbf{s}'$ is associated with three-dimensional volume elements.

We may combine the first two equations and write

$$\text{div. } (\mathbf{s}' + \mathbf{r}') = \psi,$$

and regard the two vectors as combined into one. Of this single vector the part $\mathbf{s}'$ controls the phenomena of electrodynamics, it is related to $\mathbf{F}$ by the fourth equation and through $\mathbf{F}$ to the vector potential $\mathbf{a}$.

In empty space, at any rate, we may deduce $\mathbf{G}$ from it for $\mathbf{G}$ and $\mathbf{F}$ are simply related in this case, $\mathbf{G}$ being the reciprocal of $\mathbf{F}$.

It would be a very simple scheme of things if we could refer $\mathbf{s}'$ one step farther back to some fundamental scalar ψ, and thus complete our geometrical conception of the universe by associating ψ with the four-dimensional volume.

We cannot do this directly, but we can refer to ψ a fundamental vector $(\mathbf{s}' + \mathbf{r}')$ which is such that its divergence is ψ.

This is what is meant by stating that of all these quantities ψ is the most fundamental. ψ cannot be a perfectly arbitrary scalar, it must satisfy some equation, this equation being the mathematical expression of a law of nature. We cannot hope to deduce such a law by a series of logical deductions applied to a four-dimensional world, for it will be of the same character as the principle of least action, the law of gravitation or the laws of thermodynamics, the truth of which rests upon experimental results.

Let us try the law contained in the equation

$$\mathbf{s}' + \mathbf{r}' = \kappa \text{ grad. } \psi,$$

where κ is a constant to be determined by application to special problems.

We therefore deduce the differential equation

$$\text{div. } (\mathbf{s}' + \mathbf{r}') = \psi = \kappa \text{ div. (grad. } \psi).$$

This is the relation which ψ must satisfy, and it is the fundamental law we are seeking.

The operations of divergence and gradient are to be understood in their general significance.

Grad. ψ is strictly a vector with components $\partial\psi/\partial x_m$, etc., whereas the divergence is an operator concerning a contravariant vector so that (grad. ψ) in this equation must be regarded as having for its typical component the quantity $(g^{mn}\,\partial\psi/\partial x_m)$, and the expression div. (grad. ψ) is equivalent to

$$\frac{1}{\sqrt{g}}\frac{\partial}{\partial x_m}\left(\sqrt{g}.g^{mn}\frac{\partial\psi}{\partial x_n}\right).$$

Thus the equation for ψ is

$$\frac{\kappa}{\sqrt{g}}\frac{\partial}{\partial x_m}\left(\sqrt{g}.g^{mn}\frac{\partial\psi}{\partial x_n}\right) = \psi.$$

It is significant that this is the four-dimensional mode of writing Shroedingers' wave equation.

The equation div. $\mathbf{s}' = 0$, is the equation of continuity of electricity. If we are to combine $\mathbf{s}'$ and $\mathbf{r}'$ into a single vector we must be careful that the relation

$$\text{div. } (\mathbf{s}' + \mathbf{r}') = \psi,$$

does not contradict the macroscopic equation of continuity. This appears to

impose upon ψ a microscopic variation so that ψ takes positive and negative values, and so that macroscopically

$$\overline{\text{div. } (\mathbf{s}' + \mathbf{r}')} = \bar{\psi} = 0,$$

where the bars denote average values.

Possibly this may be associated with a periodicity in time for the quantity ψ, corresponding with conclusions reached by Schroedinger. ('Ann. d. Physik,' vol. 81, p. 135 (1926).)

It may be that this is connected with the existence of point charges in space, and, if this is so, doubtless the electronic charge will enter into the relation. There will then be the possibility of linking up in this theory the fundamental electronic charge and the element of action, h, which determines the value of κ.

We must confess that we do not yet understand this point very clearly, but the theory would appear to admit of interesting enquiries in this direction.

It has been recognised since the first introduction of the Quantum Theory that a new element had to be introduced into Physics.

The point of view taken is that the four-vector $\mathbf{s}'$ is incomplete, and must be extended by the introduction of a four-vector $\mathbf{w}'$ made up of $\mathbf{s}'$ and the additional $\mathbf{r}'$, so that

$$\mathbf{w}' = \mathbf{s}' + \mathbf{r}' \ .$$

$\mathbf{r}'$ is the new element introduced.

It seems possible in this way to include in one uniform scheme gravitational, electrical and quantum phenomena.

Homogeneous Reactions involving Complex Molecules.—The Kinetics of the Decomposition of Gaseous Dimethyl Ether.

By C. N. Hinshelwood and P. J. Askey.

(Communicated by H. Hartley, F.R.S.—Received March 5, 1927.)

Chemical changes in which molecules of complex structure take part have recently assumed a particular importance in the theory of the mechanism of reaction velocity.* Decompositions of organic compounds in the gaseous state provide excellent examples of such reactions. It is often necessary, in these, to suppose the intermediate formation of unstable radicles which undergo rapid subsequent transformations; and whenever each stage of a reaction is not expressed by conventional chemical equations, involving known substances only, there may always remain a slight doubt whether the course of the reaction has been truly represented. Although this does not prevent the drawing of perfectly valid conclusions about the kinetics of a given reaction, nevertheless an example in which every stage can be represented by a simple chemical equation has undoubted advantages. The decomposition of dimethyl ether, with which this paper deals, provides such an example.

$$CH_3 . O . CH_3 \longrightarrow [CH_4 + HCHO] \longrightarrow CH_4 + H_2 + CO.$$

The complexity of the dimethyl ether molecule is only moderate, and in the light of the discussion in the papers referred to above it is interesting to note that the reaction is an almost perfect example of the transition from the unimolecular to the bimolecular type.

The remarkable influence of hydrogen on the rate of decomposition of diethyl ether is found with dimethyl ether also, and has been investigated rather more fully.

Experimental Details.—Dimethyl ether was prepared by the action of sulphuric acid on methyl alcohol, the gas evolved being absorbed in ice-cold concentrated sulphuric acid. By dropping this solution into a mixture of water and ice a convenient stream of gas was obtained, which was dried by means of quick-lime and condensed in a bulb immersed in liquid air. While the ether was solid the apparatus was thoroughly evacuated. The ether was then carefully fractionated, the middle portion being collected in gas-holders sealed to the apparatus. The gas was found to be quite pure; moreover it is to be

* 'Roy. Soc. Proc.,' A, vol. 113, p. 230 (1926); vol. 114, p. 84 (1927).

noted that concordant results for the rate of decomposition were obtained with several independent preparations.

The apparatus and method for measuring the rate of decomposition were exactly as described in earlier papers,* except that arrangements were made for working at pressures greater than one atmosphere.

The Homogeneous Nature of the Reaction.—The reaction was easily shown to be homogeneous by the usual method of comparing the rate in an empty silica bulb with the rate in the same bulb half-filled with powdered silica. Not only was there no increase in the rate of reaction in presence of the powder, but an apparent slight decrease.

Initial pressure of dimethyl ether.	Time for 50 per cent. change.	
	′ ″	
179	5 23	Empty bulb.
175	6 13	Bulb + powdered silica.
118	5 45	Empty bulb.
119	6 40	Bulb + powdered silica.

The homogeneity of the reaction, moreover, made itself evident all through the experiments in the exact reproducibility of results, and in the nature of the kinetic relationships.

The Products of the Reaction.—In the region of 500° C., dimethyl ether decomposes quantitatively into carbon monoxide, methane and hydrogen, the reaction taking place with an increase of pressure equal to twice the initial pressure. The following is an analysis of the products of decomposition at 555° C.

	Per cent.
Carbon monoxide	32·0
Hydrogen	33·5
Methane.....	34·5

This corresponds to the equation $CH_3 . O . CH_3 = CO + H_2 + CH_4$. No traces of ethylene or water are formed.

The increase of pressure was nearly equal to the theoretical value.

Initial pressure of ether (mm.)	32	56·5	59	113	313	420
Increase of pressure (mm.)	63	113	117	225	619	838
Per cent. of theoretical increase.....	98·4	100	99·2	99·6	99·2	99·7

* 'J. Chem. Soc.,' vol. 125, p. 396 (1924); 'Roy. Soc. Proc.,' A, vol. 111, pp. 245, 380 (1926).

Observation of the pressure increase provides a convenient means of following the course of the reaction. Since, however, the change may take place in consecutive stages, it is necessary to decide how far at any moment the pressure increase is proportional to the actual amount of dimethyl ether which has decomposed. Accordingly, samples of gas were withdrawn from the reaction vessel when the pressure increase indicated an apparent decomposition of 25 per cent. and of 50 per cent. These were analysed by absorbing the unchanged ether in cold concentrated sulphuric acid and determining the remaining gases in the usual way.

When the apparent decomposition amounted to 53 per cent. the analysis was as follows :—

	Per cent.	
	Calculated.	Observed.
Unchanged ether	22·8	21·0
Carbon monoxide	25·7	22·5
Hydrogen	25·7	25·8
Methane.....	25·7	30·7

When the apparent decomposition was 27 per cent. the results were :—

	Per cent.	
	Calculated.	Observed.
Unchanged ether	47·4	44·4
Carbon monoxide	17·5	15·4
Hydrogen	17·5	18·2
Methane.....	17·5	22·0

It appears then that the rate of increase of pressure coincides nearly, but not exactly, with the actual rate of disappearance of ether. For most practical purposes it can be used to measure the velocity of reaction.

An inspection of possible equations for the decomposition shows that formaldehyde may be an intermediate product. Although it is known from the work of Bone and Smith* that the rate of decomposition of formaldehyde is rapid even at 400° C., it seemed possible that small quantities might accumulate during some part of the reaction. Its presence was detected by the colorimetric method of Schryver.† As this method did not prove suitable for the quantitative estimation, the ordinary iodimetric method was adopted, samples of the gas at various stages of the decomposition being withdrawn and shaken

* ' J. Chem. Soc.,' vol. 87, p. 910 (1905).

† ' Roy. Soc. Proc.,' B, vol. 82, p. 226 (1910).

with water. Blank tests with ether solutions showed that they were without effect on iodine. The detailed results of these analyses appear in a later section, but it may be pointed out here that the accumulation of small amounts of formaldehyde accounts for the discrepancies in the analyses recorded above. For example, when dimethyl ether at an initial pressure of 405 mm. had decomposed apparently to the extent of 25 per cent., 50 mm. formaldehyde were found to be present in the reaction vessel. If the actual amount of ether decomposed is x then $(405 - x) + 50 + x + (2x - 50) = 405 + 202{\cdot}5$ whence $x = 126$. An apparent decomposition of 25 per cent. thus corresponds really to 31 per cent. Similarly when the apparent decomposition was 50 per cent. the true decomposition was 55 per cent.

The analysis of the products of reaction withdrawn at an apparent decomposition of 53 per cent., given above, may now be compared with the following calculated analysis, in which the formation of formaldehyde is allowed for. The formaldehyde was assumed all to have condensed out as paraformaldehyde before the analysis was carried out.

	Per cent.
Unchanged ether	21·5
Carbon monoxide	24·3
Hydrogen	24·3
Methane	29·8

This is now in satisfactory agreement with the observed values.

The apparent percentage decomposition requires a correction of a few per cent. to bring it to the true value. Since it is shown in a later section that the amount of formaldehyde produced at any stage of the reaction is strictly proportional to the initial pressure of the ether, this correction is the same for all initial pressures, a fact which is important because it is often theoretically convenient to express results in terms of the time taken for a definite fraction of the total ether to react, though what particular fraction is chosen may not matter. Thus if the course of the reaction is being followed by pressure measurements, an apparent decomposition of 25 per cent. can be used without correction for comparative purposes; but may be reduced to the true value if necessary by applying the same correction whatever the initial pressure may be.

The Kinetics of the Reaction.—When the increase of pressure, x, is plotted against the time, t, curves approximating closely to the unimolecular form are obtained. In the tables below T is the Centigrade temperature, a the initial

pressure of dimethyl ether; k is the "uncorrected" unimolecular constant obtained from the expression $k = \frac{1}{t} \log_{10} \frac{2a}{2a - x}$, the total pressure increase being equal to $2a$.

T = 504°; $a = 312$. T = 552°; $a = 420$.

t.	x.	$k \times 10^3$.	t.	x.	$k \times 10^2$.
′ ″			′ ″		
3 27	56	0·198	0 57	164	0·166
6 30	96	0·186	1 25	242	0·174
8 1	116	0·186	1 54	323	0·185
11 5	156	0·188	2 25	395	0·190
12 57	176	0·185	3 2	471	0·196
15 16	200	0·183	3 39	534	0·202
19 55	250	0·186	4 21	593	0·204
26 27	312	0·190	4 59	634	0·204
37 20	402	0·200	5 45	675	0·205
44 20	437	0·197	7 44	744	0·203
52 35	467	0·190	9 24	778	0·204
∞	619		∞	838	

T = 510°; $a = 59$.

t.	x.	$k \times 10^4$.
′ ″		
3 10	6	0·119
11 0	20	0·122
22 20	40	0·134
28 30	50	0·140
35 5	59	0·143
43 50	70	0·148
54 40	80	0·150
67 0	88·5	0·146
101 0	103	0·148
∞	117	

The rise in the constants in the early part of the reaction may be attributed to the existence of the consecutive reactions, the second of which does not become as fast as the primary decomposition until a small amount of formaldehyde has accumulated. This point is soon reached, but until it is the ether decomposition does not produce the full pressure increase per molecule.

Nevertheless the shape of the curves corresponds quite closely to a reaction of the first order. On the other hand the values of k, as will appear later, are not independent of the initial pressure when this is below about 400 mm.; that is to say, the reaction is not strictly unimolecular at lower pressures—as inferred

from the influence of initial pressure the order approaches more nearly to the second. Yet the shape of the curves connecting t and x remains the same at all pressures, a curious point which is fully explained later on. The following table shows that although the times required for x to reach 25 per cent., 50 per cent. and 75 per cent. respectively of its final value vary with the initial pressure, the ratio of these times among themselves is the same whatever the initial pressure may be.

Initial pressure of dimethyl ether.	T.	Actual values.			Relative values. $t_{25} = 1{\cdot}0$.		
		t_{25}	t_{50}	t_{75}	t_{25}	t_{50}	t_{75}
	°C.	′ ″	′ ″	′ ″			
32	555	4 13	8 55	16 50	0·473	1	1·89
41	555	3 50	7 40	15 5	0·500	1	1·97
56·5	510	16 30	35 30	70 0	0·47	1	1·97
59	510	16 20	35 0	68 0	0·47	1	1·94
91	504	19 40	43 0	88 0	0·46	1	2·05
150	504	15 0	37 0	77 0	0·41	1	2·08
171	504	14 0	31 30	62 30	0·44	1	1·98
233	555	1 38	3 20	6 30	0·49	1	1·95
312	504	11 24	26 27	53 20	0·43	1	2·01
420	552	1 14	2 37	4 55	0·47	1	1·88
518	552	1 13	2 38	5 17	0·47	1	2·01

It will be seen that the shape of the t, x curves is the same at all temperatures and pressures. The unimportance of the variation among the figures for the relative times is shown by the fact that a change in the shape of the curve from a unimolecular to a bimolecular form would cause the ratio t_{75}/t_{50} to change from 2·0 to 3·0.

From the similarity in shape of the different curves it would be inferred that any effect due to the production of formaldehyde must be proportional to the initial pressure. This is confirmed experimentally.

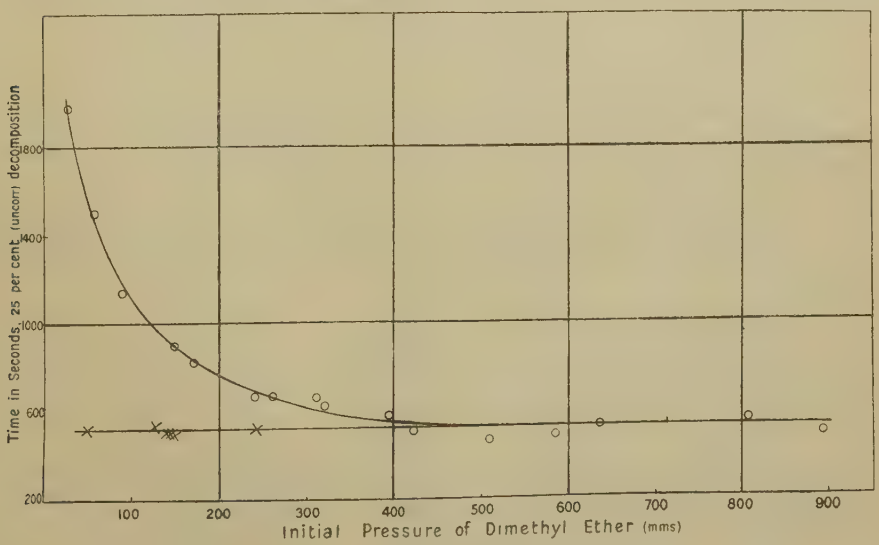
Initial pressure of dimethyl ether.	Apparent percentage decomposition.	Actual Pressure of formaldehyde.
mm.		mm.
401	5	31
412	10	39
405	25	50
396	50	43
416	75	22
112	10	11
115	25	13
113	50	9
123	75	6

Comparison of the two halves of the above table shows that the amount of formaldehyde present at any stage of the reaction is proportional to the initial pressure of the dimethyl ether.

Influence of the Initial Pressure.—At high pressures the reaction obeys the unimolecular law, the time taken for any given fraction of the ether to decompose being independent of the initial pressure. Below about 400 mm., however, this time increases and the reaction assumes more and more a bimolecular character. This may be illustrated conveniently by recording the times required for an apparent decomposition at 25 per cent. at 504° C. (*i.e.*, a true decomposition of 31 per cent.).

Initial pressure.	Time.	Initial pressure.	Time.
mm.	′ ″	mm.	′ ″
28	33 0	321	10 25
58	25 0	394	9 50
91	19 0	422	8 28
150	15 0	509	7 45
171	13 44	586	8 4
241	11 7	636	8 58
261	11 10	807	9 19
312	11 5	894	8 17

These values are plotted in the figure. The plotting of the times required for



any other fraction gives similar results. The special investigations described in the previous paragraph show that these figures cannot be influenced by the

intermediate formaldehyde formation; and thus they represent truly the effect of the initial pressure of the ether.

Influence of Hydrogen on the Rate of Reaction.—In presence of a sufficient quantity of hydrogen the reaction preserves its unimolecular character down to low pressures as shown by the lower curve in the figure.

Initial pressure of dimethyl ether.	Initial pressure of hydrogen.	Time for apparent 25 per cent. decomposition.
		′ ″
27	400	8 58
50	403	7 56
140	638	8 21
146	672	8 21
150	401	8 15
150	600	7 55
244	397	8 43

The behaviour in presence of increasing quantities of hydrogen when the initial pressure of the ether is low is very significant. The rate of reaction increases at first until the hydrogen pressure is about 300 mm. At this point the rate attains the value characteristic of high partial pressures of ether and cannot be caused to increase any more however much hydrogen is added. Nor has hydrogen any effect when the initial pressure of ether is high enough to bring the reaction within the " unimolecular " region. This all seems to disprove any suggestion that the effect of the hydrogen depends upon the formation of intermediate compounds, " complexes " or reduction products. It is difficult to escape from the conclusion that the effect is in some way connected with the maintenance of enough activating collisions, which in the absence of hydrogen would begin to be insufficient at about 400 mm. of dimethyl ether.

The following series of experiments made at 504° C. may be quoted in illustration:—

Initial pressure of dimethyl ether.	Pressure of hydrogen.	t_{25}.	t_{50}.	t_{75}.
		′ ″	′ ″	′ ″
150	0	15 0	36 45	77 0
154	200	11 5	26 0	52 35
154	300	8 43	22 35	—
148	400	7 41	21 20	—
150	600	7 55	17 55	36 30
Compare limiting values for high initial pressure of ether		8 0	17 23	34 46

It may be mentioned that estimations of formaldehyde were made in experiments in which the decomposition took place in presence of hydrogen. The results were normal, showing that the reducing action of the hydrogen under these conditions is negligible.

The influence of hydrogen accounts for the curious fact that the velocity constants are only independent of the initial pressure of the ether when this is above 400 mm., and yet remain constant throughout the course of a decomposition. Hydrogen is one of the products of reaction; it will be seen from the table above that, for a low pressure of ether of 150 mm., the limiting rate of reaction is reached when the pressure of hydrogen is about 300 mm. The limiting rate is also reached with about 400 mm. of dimethyl ether. Thus dimethyl ether and hydrogen are about equally effective in keeping up the number of activating collisions. Since, moreover, one molecule of ether yields one of hydrogen in decomposing, the specific reaction rate can remain nearly constant throughout the course of the reaction. In this way it is possible to explain the apparent paradox of unimolecular velocity constants which vary with the initial pressure.

Influence of other Gases.—Neither helium, nitrogen, carbon monoxide, carbon dioxide nor methane have an effect similar to that of hydrogen.

Helium has no influence—

T.	Initial pressure of dimethyl ether.	Pressure of helium.	t_{25}.	t_{50}.	t_{75}.
°C			′ ″	′ ″	′ ″
504	116	200	16 52	39 5	78 0
504	116	0	17 20	38 30	77 0

Carbon monoxide has, if anything, a slight retarding effect—

T.	Pressure of dimethyl ether.	Pressure of carbon monoxide.	t_{25}.	t_{25} in absence of CO.
°C			′ ″	′ ″
504	108	398	19 31	17 50

Carbon dioxide also appears to have a slight retarding effect if present at considerable pressures.

Nitrogen had little influence; in an experiment with 235 mm. of dimethyl ether and 400 mm. of nitrogen, where the total pressure increase was 460 mm.

out of the theoretical 470, t_{25} was 12′ 45″. The corresponding value in the absence of nitrogen would have been 11′ 40″.*

The Temperature Coefficient and Heat of Activation.—Experiments to determine the temperature coefficient of the reaction velocity were made with pressures of dimethyl ether sufficiently high to ensure that the values found for the rate were the limiting values corresponding to the "unimolecular" region.

At least two experiments were made at each of six temperatures, and curves were plotted from which the average value of k and the value t_{25} were found.

The results are summarised below, k being expressed with the time in seconds, and with natural logarithms.

T (abs.).	t_{25}.		k (average).	k (calculated).
°	′ ″	″		
825	1 13 1 13	73	0·00431	0·00430
795	4 17 4 0	248	0·00113	0·00112
777	8 28 8 0	494	0·000544	0·000467
751	38 0 38 26	2,293	0·000127	0·000125
725	154 0 144 0	8,940	0·0000317	0·0000311
695	850 0 865 0	51,420	0·00000447	0·00000531

The value of E calculated from the Arrhenius equation is 58,000 calories from the values of t_{25}, and 58,500 calories from the values of k.

$k_{\text{calculated}}$ is obtained from the equation

$$\ln k = 30{\cdot}36 - \frac{58{,}500}{RT}.$$

At 800° (abs.) the value of k is 0·00140 in the steady region, and it begins to fall at about 400 mm. At this temperature and pressure the number of mole-

* The first experiments made in the presence of nitrogen gave rather curious results, the reaction failing to reach its proper end-point. At first, it seemed as though some different reactions were taking place, but analysis of the products showed that carbon monoxide, methane and hydrogen were present in exactly the normal proportions, and there was neither unchanged ether, nor any solid or condensable product. The use of fresh samples of nitrogen, carefully freed from oxygen, and careful manipulation removed the anomaly.

cules, N, which react per cubic centimetre in unit time is therefore $6{\cdot}8 \times 10^{15}$. Taking the molecular diameter as 5×10^{-8} cm., the total number of collisions, Z, at 800° abs. is $1{\cdot}73 \times 10^{28}$. Assuming that at 400 mm. the rate of activation, calculated in the manner described in earlier papers, is equal to the rate of reaction

$$N = \frac{Z \cdot e^{-\frac{\{58{,}500 + (\frac{1}{2}n - 1)RT\}}{RT}} \left(\frac{58{,}500 + (\frac{1}{2}n - 1)\,RT}{RT}\right)^{(\frac{1}{2}n - 1)}}{\lfloor \frac{1}{2}n - 1},$$

where n is the number of energy terms among which the energy of activation is distributed.

Introducing the values of N and Z

$$\frac{e^{-(\frac{1}{2}n-1)}\{36{\cdot}93 + (\frac{1}{2}n - 1)\}^{\frac{1}{2}n-1}}{\lfloor \frac{1}{2}n - 1} = \frac{6{\cdot}8 \times 10^{15}}{1{\cdot}73 \times 10^{28}} \times e^{36{\cdot}93} = 4300.$$

If $n = 10$, the left-hand side is 2140, while if $n = 12$ it is 7270. n may therefore be taken as 11.

The value of n found for diethyl ether is 8 and for propionic aldehyde about 12.

Mr. Fowler considers it not impossible that under certain conditions all the energy of two colliding molecules may pass into one. He takes n, calculated as above, to be the sum of the energy terms of the two molecules in collision. Although we still feel that this is a very extreme assumption, something like it seems to be necessary to account for the rate of activation of nitrogen pentoxide molecules ; and a value of 11 for n is perhaps rather a large one for a single molecule of dimethyl ether. On the other hand the methyl group is disrupted, so that the vibrational degrees of freedom concerned in the attachment of the six hydrogens to their carbon atoms are probably all in play. Thus a value of n equal to 11 is by no means impossible.

Discussion of the Influence of Inert Gases.—No complete explanation is available of the remarkable difference between hydrogen and other gases. It has already been pointed out that hydrogen in all probability acts only by maintaining the Maxwell distribution among the molecules of dimethyl ether when this would otherwise begin to be disturbed. The ground for believing this is that hydrogen cannot do more than restore the rate of reaction to its normal limiting value. The rather surprising fact is that other gases do not have a similar effect. Hydrogen has a greater molecular velocity than any other gas, nearly four times greater than nitrogen. Thus there will be four times as many collisions, approximately, between ether and hydrogen as

there would be between ether and nitrogen at the same pressure. Helium, which is the only gas with a molecular velocity approaching that of hydrogen, has only three degrees of freedom and therefore can communicate much less energy than the diatomic gases. Qualitatively, therefore, the unique position of hydrogen is understandable, but quantitatively the differences seem greater than can be accounted for in this way. The advantage that hydrogen possesses, in virtue of its great velocity over all gases except helium, and over helium in virtue of its five degrees of freedom, seems to be reinforced by some specific factor. It is as though, for example, molecules of ether could store up energy from successive hydrogen collisions, in a way which is not possible with nitrogen and the other gases.

Summary.—The decomposition of dimethyl ether takes place according to the equation $CH_3 . O . CH_3 \longrightarrow CH_4 + HCHO \longrightarrow CH_4 + CO + H_2$, the formation of formaldehyde being but transitory.

The reaction is empirically unimolecular at pressures above 300–400 mm., but at lower pressures ceases to be independent of the initial pressure. In the presence of sufficient hydrogen the reaction maintains its unimolecular character at all pressures. The hydrogen seems to act only by preserving the Maxwell distribution among the molecules of the dimethyl ether when this would otherwise be disturbed by the chemical transformation of activated molecules ; for it cannot do more than restore the rate of reaction to its normal limiting value, and cannot increase it beyond this. Nitrogen, helium, carbon dioxide and carbon monoxide do not have a similar influence.

The limiting rate of reaction is expressed by the equation

$$ln\, k = 30{\cdot}36 - \frac{58{,}500}{RT}.$$

To account for the observed rate of reaction, the energy of activation must be supposed to be distributed between about 11 energy terms.

An Experimental Test of the Dipole Theory of Adsorption.

By W. G. Palmer, Fellow of St. John's College, Cambridge.

(Communicated by Sir William Pope, F.R.S.—Received March 9, 1927.)

In previous communications* an account was given of the results obtained when the phenomenon of electric coherence is applied to the investigation of adsorption on plane metallic surfaces. The nett heat of desorption per unit area of surface is given by

$$Q_a = [E^2(K - 1)/d] \times 0{\cdot}53 \times 10^{-5} \text{ calories}, \qquad (1)$$

where E is the critical or cohering voltage, K is the specific inductive capacity of the film, and $2d$ is its thickness. For a film completely covering the surface the molecular heat of desorption is

$$Q_m = [AE^2(K - 1)/d] \times 16{\cdot}0. \qquad (2)$$

A being the area occupied on the surface by one molecule ; in the above expressions the unit of d is 10^{-8} cm., and of A, 10^{-16} cm. The total heat liberated on adsorption of 1 gram-mol. is $Q_m + L$, where L is approximately the latent heat of evaporation per gram-mol. In short, Q gives a direct measure of the maximum adsorption potential, as this is defined by Polanyi and by Lorenz and Landé.†

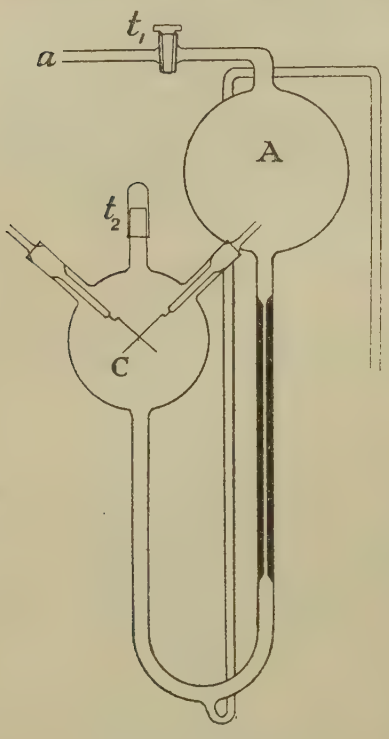
The method formerly described has now been adapted without essential modification to the study of adsorption from liquids, in which the cohering junction is completely immersed. Owing to the appreciable conductivity of liquids other than hydrocarbons, the condenser formerly used as a means of applying electric tension to the junction has in the present work been replaced by direct connection to the potentiometer through 10,000 ohms resistance. A comparison of the two arrangements with hydrocarbons and the practically non-conducting higher alcohols showed that both methods gave identical results.

A diagram of the coherer adapted for liquids is shown in the figure. The mode of introducing liquid is as follows : the tap t_1 is opened, while t_2 is kept closed, and liquid is drawn into A by a pump attached at a. The pump is disconnected, and t_2 opened, when liquid passes from A through the capillary into the coherer C. The capillary tube reduces the rate of flow into C so that the filaments are

* 'Roy. Soc. Proc.,' A, vol. 106, p. 55 (1924) ; vol. 110, p. 134 (1926).

† Lorenz and Landé, 'Z. f. anorg. Chem.,' vol. 125, p. 48 (1922).

not deranged. The latter are brought into mechanical contact by turning in ground glass joints as in the previous experiments. Throughout the present work platinum filaments were used; these were cleaned by ignition in an alcohol flame before experiments on a series of liquids were undertaken, but were not subsequently withdrawn from the coherer until the results for the whole series were completed, thus ensuring that the surface has substantially the same condition throughout. All the experiments in the present work were carried out at 18 to 20° C.



Primary Alcohols.

Methyl alcohol was obtained from the commercially pure liquid by conversion into oxalate, which was hydrolysed after recrystallisation and the product rectified. The other alcohols were obtained from the commercial pure material by fractionation to within 0·2° C. All the samples were dried by digestion with fresh lime immediately before they were to be used.

The values tabulated under l represent (in units of the diameter of the carbon atom) the effective length of the alcohol chain, the length of the group $-CH_2OH$ being taken as twice the diameter of the carbon atom, since the diameter of the oxygen atom is nearly equal to that of carbon. The values of the critical voltage E are, if necessary, corrected for conductivity; this was appreciable only with methyl alcohol and formic acid, but all traces of conductivity even in these cases disappeared when the liquid was lowered away from the filaments, although these were separated by the film only; the conductivity does not, therefore, take place through the film but between the other parts of the filaments and particularly the relatively large leads.

Table I.

Alcohol.	Critical voltage.	l.	E^2/l.
Methyl	3·0	2	4·5
Ethyl	3·7	3	4·56
Normal propyl	4·25	4	4·5
Normal butyl	4·7	5	4·4
Iso-butyl	3·8	(4·5)	(3·2)

The Fatty Acids.

The samples of formic and acetic acids melted at 7·4° C. and 15·8° respectively. Propionic acid was prepared by the hydrolysis of the pure nitrile, and butyric acid by the usual malonate synthesis. Heptoic acid was obtained by the oxidation of pure œnanthol. In the table under l the length of the group —COOH is taken as equal to that of $-CH_2OH$.

Table II.

Acid.	Critical voltage.	l.	E^2/l.
Formic	2·6	2	3·4
Acetic	3·6	3	4·3
Propionic	4·0	4	4·0
Butyric	4·5	5	4·0
Heptoic	6·5	8	5·25

Ethyl Esters of the Fatty Acids.

These were obtained by esterifying the pure acids. The length of the group $-COOC_2H_5$ is taken as four times that of the carbon atom.

Table III.

Ester.	Critical voltage.	l.	E^2/l.
Formic	5·2	4	7·0
Acetic	6·25	5	7·8
Butyric	6·75	7	6·5
Heptoic	8·75	10	7·6

The values of the voltage for these esters could only be obtained to an accuracy of about 0·3 volt, owing probably to the great difficulty in preserving an ester free from traces of acid in glass vessels. The values recorded above are the means of several observations.

Hydrocarbons.

(*a*) A series of paraffin fractions were prepared from petrols of varying boiling points. Their approximate compositions were found from their boiling range and the specific gravity.

	Boiling point.	Composition by volume.
C_5 fraction	40–50°	*n*-Pentane and cyclopentane in equal amounts.
C_6 fraction	67–77°	Methyl cyclopentane 25 per cent. Cyclohexane 25 per cent. *n*-Hexane 50 per cent.
C_7 fraction	100–115°	Methyl cyclohexane and *n*-heptane in equal amounts.

(*b*) Samples of cyclohexane and cyclohexene from Messrs. Poulenc were found to be pure. Benzene was freed from sulphur compounds by heating with aluminium chloride, and then further purified by freezing. In the case of the C_5 fraction above, it was possible to compare the voltage in the saturated vapour with that in the liquid, since the vapour pressure at 20° C. was 461 mm., and it was easy to arrange that this pressure caused such a depression in the level of the liquid in the bulb of the coherer that the filaments were exposed to vapour only. The following results were obtained :—

C_5 (pentane) fraction—

	Voltage.
Filaments in liquid	0·3
Liquid lowered ; filaments in vapour	0·3
Liquid again raised	0·4
After 24 hours in liquid	0·4

It is thus clear that the adsorbed film has the same properties whether it be produced from the saturated vapour or from contact with actual liquid. The results of this experiment confirm the statement (*vide supra*) that the electrical method measures the nett heat of adsorption.

The voltage for the paraffin fractions increases somewhat with time of contact, indicating probably an increase on the surface of the amount of the cyclic constituent.*

* ' Roy. Soc. Proc.,' A, vol. 110, p. 134 (1926).

Table IV.

Hydrocarbon.	Critical voltage.	
C_5 fraction	0·4	
C_6 fraction	0·6	
After three hours	0·9	(*cf.* cyclohexane below).
C_7 fraction	0·9	
After one hour	1·1	
Cyclohexane	0·8	
Cyclohexene	2·5	
Benzene	0·9	

The voltages of the last three pure liquids were unaffected by time of contact.

The Mechanism of Adsorption.

Debye* and others have shown that many quantities expressing the physical inter-relations of molecules can be successfully explained and calculated on the assumption that molecular attraction is caused by the interlocking of electrically polar structures, and it seems reasonable to seek the origin of adsorptive forces in the same direction. Lorenz and Landé† have already discussed adsorption upon a plane metallic surface from this standpoint, and have overcome the objections raised by Polanyi.‡ They assume that a molecule is attracted to and subsequently held on the surface by the electrical image of its dipole in the surface ; and they arrive at the following conclusions (for a substance with a single dipole)—

(*a*) the adsorption layer is not more than one molecule in thickness ;

(*b*) the nett heat of adsorption is given by

$$Q_m = RM^2/4x^3k \qquad (3)$$

where R is the gas constant, M the electric moment of the adsorbed molecule, x the distance of the centre of the dipole axis from the surface and k is Boltzmann's constant. This expression for Q has a much wider range of validity than Lorenz and Landé suggest. An inspection of the series from which it is derived shows that it holds with almost no loss of exactness up to $x = l$, l being the length of the dipole axis ; it will be rare for the dipole in a molecule to approach the adsorbing surface so close as this limit implies.

* 'Phys. Z.,' vol. 21, p. 180 (1920).

† *Loc. cit. supra.*

‡ 'Z. f. Elektrochemie,' vol. 26, p. 370.

Using a known relation between the specific inductive capacity and the moment M,* we find for a polar phase to which Maxwell's law applies

$$K - 1 = [0{\cdot}88M^2/T] \,.\, Df/Dg \times 10^{36} \qquad (4)$$

where T is the temperature, Df is the density of the molecules in the phase considered, and Dg is the density in the vapour supposed at 0° C. and 760 mm.

In the alcohols, acids and esters, respectively, the principal doublet is undoubtedly in the groups CH_2OH, COOH and $COOC_2H_5$, and therefore the adsorbed molecules will, in our view, be attached to the surface by these groups. We shall have, in fact, the same sort of orientation in the adsorbed film as that found by Langmuir, Adam and others on the surface of water. Hence in the above expression for K, Df remains unchanged within a homologous series. For the same reason, x in (3) above also is constant in a homologous series. Hence the dipole theory predicts that in a given series the heat of desorption Q_m will be simply proportional to the moment of the adsorbed molecules, and this by (4) is itself proportional to K — 1 for a fixed temperature. We therefore expect to find $Q_m = k\,(K - 1)$, where k will vary with the series.

The third column in the tables above shows that E^2/l is closely constant for the alcohols, and very fairly so for the other two series. Referring to (2), and noticing that l must be proportional to the thickness d, this prediction is seen to be very closely confirmed.

Since both the expression (3) and the relation (4) depend on the applicability of Maxwell's distribution law to the substance concerned, we must conclude that orientation implies only attachment to the surface by one group, and not a permanent alignment of the adsorbed molecules. Apart from the improbability of the existence of completely oriented or "solid" films of alcohol at room temperature, a confirmation of the type of film can be obtained in another way. If the molecules in the film are all arranged as they would be in solid alcohol, then in calculating the heat of desorption from (2) we must take K as the value for the solid, and this differs but little from the square of the refractive index of the liquid. Putting $K = 2{\cdot}5$, taking $A = 21{\cdot}7$ and $d = 1{\cdot}3\,l$,† we find for a solid alcohol film $Q_m = 1900$ calories. On the other hand, if the film is liquid, as is suggested above, then K is obtained from the known values of M.‡ This gives $Q_m = 7000$ calories. Now Q_m cannot be very different in magnitude from the heat of association of the liquid concerned. Unfortunately the heat of association does not appear to have been found for alcohol, but since

* J. J. Thomson, 'Phil. Mag.,' vol. 27, p. 763 (1914).

† Adam, 'Roy. Soc. Proc.,' A, vol. 101, p. 463 (1922).

‡ Smyth, 'J. Amer. Chem. Soc.,' vol. 46, p. 2155; vol. 47, p. 1894.

the degree of association of water and ethyl alcohol are nearly the same at the same temperature, we may assume that the heats of association are of the same order; for water the heat is 9640 calories,* and this is much nearer to the second than to the first value of Q above.

From the equation $Q_m = RM^2/4x^3k = 7000$ we obtain $x = 1{\cdot}3 \times 10^{-8}$ cm., thus confirming the view that the molecule is attached to the surface by the polar group.

It appears from the work of Smyth† that the values of M are nearly constant for the same homologous series; in the alcohols 1·9–2·0, fatty acids 1·3, and ethyl esters 1·5, but only for the alcohols is the value well-established. It may therefore be inferred that the heat of desorption from a platinum surface of all the primary alcohols is approximately 7000 calories; for the acids and esters the value is probably somewhat less. Gaudechon‡ found almost constant values for the heat of wetting of both clay and silica by the primary alcohols and fatty acids respectively; his method allows only the heat per gram of adsorbent to be calculated.

It will be noticed that in Table I isobutyl alcohol shows an apparent marked divergence from the rule $E^2/l = 4{\cdot}5$. This alcohol has the formula $(CH_3)_2CH \cdot CH_2OH$ and therefore contains a branched chain. The molecule will have a greater cross-section than the simple normal chain alcohols and at the same time less length. If we take the latter as intermediate between those of *n*-propyl and *n*-butyl alcohols, namely as 4·5 units, we find $E^2/l = 3{\cdot}2$. As there is no reason to suppose that Q_m in this case differs from 7000 calories, we must have $A_{isobutyl}/A_{normal} = 4{\cdot}5/3{\cdot}2$ (see equation (2)). Thus the branched chain occupies about 1·4 times the area taken on the surface by a normal chain, and this is in harmony with a simple tetrahedral model of the branched chain.

The Mechanism of Adsorption of Non-polar Molecules.

The dipole theory seems to account so well for the experimental facts in the case of the permanently polar substances used in the work described above, that it may be worth while to generalise it. We may assume that molecules which give rise to normal, non-associated liquids, and have therefore weak external fields, will not be capable of adsorption in their normal state, at least on metallic surfaces. However, if two (practically) non-polar molecules undergo thermal collision of sufficient violence, moments of considerable magnitude may

* W. E. S. Turner, 'Molecular Association,' p. 17.

† *Loc. cit.*

‡ 'Comptes Rendus,' vol. 157, p. 209.

be produced but will normally exist only for a very short time. If such an "activating" collision occurred near a metallic surface, the existence of the temporary dipole would be prolonged by the immediate attraction of the electric image in the surface; and the activated molecule may as a result be held on the surface, but only while it remains activated. It will be noticed that the energy of activation is the whole energy associated with the temporary doublet, while the energy of adsorption represents the energy liberated by the extinction of only that part of the molecular field that is external to the molecule. In general therefore there will be no quantitative relation between the energy of activation of the adsorbed molecule and the heat developed on adsorption, except that the latter will be the smaller quantity. The many recent unsuccessful attempts to correlate heat of adsorption with heat of activation in connection with surface catalysis only serve to emphasise this conclusion. Without going deeply at the present juncture into the bearing of the above suggestions on the problem of surface catalysis, it may be remarked that the existence of catalytic activity proves that some at least of the adsorbed molecules are highly activated, and it has never been easy to reconcile a positive heat of adsorption, sometimes of considerable magnitude, with the endothermic process of activation; in fact it might be expected, in the absence of any special theory, that an adsorbed molecule would be more and not less stable than a free molecule.

It may be objected that the theory demands that the amount of adsorbed substance should be directly proportional to the external pressure, whereas it is known to vary in general more slowly. It is true that the rate of formation of the film would be proportional to the pressure, but if, as it is reasonable to suppose, evaporation is a consequence of collision with molecules from outside the layer, then the rate of evaporation will also be proportional to the pressure, and the amount in the film may actually vary only slowly with change of pressure.

Adsorption of Hydrocarbons on Platinum (*Table* IV).

If E be plotted against the number of carbon atoms in the hydrocarbon chain, it is found that the cohering voltage will be nearly zero for a 4-carbon chain. It may hence be inferred that pentane is the lowest hydrocarbon of the series that can produce an adsorption film on platinum. As has been explained above, it appears that in these cases of non-polar bodies we are concerned with the possibility of the production of temporary polar molecules before adsorption can occur. It will be in accordance with the general conception of a paraffin chain to suppose that any energy acquired by collision is shared equally by

all the carbon atoms in the chain. At the same time the number of degrees of freedom in the whole molecule increases (possibly linearly) with the number of carbon atoms, and therefore in accordance with the views of Lindemann and Hinshelwood and others,* the energy that can be acquired in similar collisions by an n-C hydrocarbon is greater than that acquired by an $(n - 1)$ C hydrocarbon. Hence we may suppose that in each carbon atom in a paraffin chain there may be created in the same conditions approximately equal moments. The experiments show that at least five of these moments are necessary to hold the chain in attachment to the surface in opposition to the thermal movements.

It will follow, since a hydrocarbon is an elongated structure of some rigidity, that hydrocarbons will be adsorbed with the length of the chain parallel to the surface; and ring hydrocarbons such as the cyclo-paraffins with the plane of the ring parallel to the surface. It is of interest to notice that benzene and cyclohexane behave so similarly; the unsaturated (ethylenic) cyclo-hexene, however, has a much greater cohering voltage, and approaches the values for permanently polar substances. At the same time in view of its physical properties there can be no doubt that this hydrocarbon is only very weakly polar at most. It will only be stating the chemical properties of ethylenic compounds in another way to say that in such substances energy of activation can remain localised in the ethylenic bond. We may therefore foresee that cyclo-hexene can remain adsorbed even when only the ethylenic bond is in contact with the surface, and the ring is, so to speak, on edge. There will hence be two points of difference between the films of cyclo-hexene and of cyclo-hexane, (a) the film of the former will contain more molecules in unit area, and (b) it will be thicker in the ratio $\frac{\text{cross-section of the ring}}{\text{thickness of the ring}}$. These considerations account for the high voltage of cyclo-hexene.

It is well known that when benzene is catalytically hydrogenated on the surface of nickel, only the completely reduced hexahydrobenzene is produced, unaccompanied by any intermediate products; nor are any intermediate products obtained when cyclo-hexane is dehydrogenated by the action of platinum or palladium. It is very difficult to explain these observations if the hydrocarbons could be attached to the surface by an edge alone, but it is an obvious consequence of attachment by all the atoms in the ring.

* Hinshelwood, 'Roy. Soc. Proc.,' A, vol. 113, p. 230 (1926).

Summary.

(1) The electric coherer functions normally when the loose contact is immersed in liquids.

(2) The cohering voltage increases regularly in the homologous series of primary alcohols, fatty acids, and their ethyl esters according to the rule E^2/l = constant, where l is the length of the chain.

(3) This experimental result is shown to indicate that the energy of desorption in a given series is proportional to the square of the electric moment of the adsorbed molecule, and it thus strongly supports the dipole theory of adsorption.

(4) An attempt is made to extend the theory to include the adsorption of substances such as hydrocarbons that are not normally more than very weakly polar.

The Thermal and Electrical Conductivity of a Single Crystal of Aluminium.

By Ezer Griffiths, D.Sc., F.R.S., Physics Department, National Physical Laboratory.

(Received March 3, 1927.)

The recent work of Carpenter and Elam on the growth of single crystals of large dimensions has rendered possible the study of the physical constants of single crystals of the commoner metals, and the present communication describes the determination of the thermal and electrical conductivity of aluminium in the form of an isolated crystal.*

The form of the crystal investigated is shown in fig. 1. This crystal had been prepared at the National Physical Laboratory employing the technique

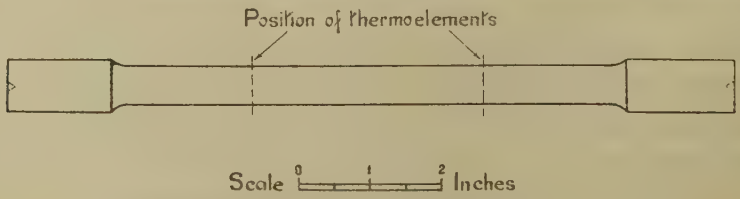


Fig. 1.—Single Crystal of Aluminium.

* Mr. Gough, M.B.E., Ph.D., of the Engineering Department, kindly supplied the crystal.

described by Carpenter in "Nature," p. 266, August 21, 1926, which briefly is as follows :—

The test specimen is machined and subjected to three treatments, thermal, mechanical, and thermal. The first treatment is necessary to soften the metal completely and produce new equiaxed crystals of so far as possible uniform size, the average diameter being 1/150 inch. The second consists in straining these crystals to the required amount, and the third in heating the strained crystals to the requisite temperature, so that the potentiality of growth conferred by strain could be brought fully into operation.

X-ray analysis* of the specimens by the rotating crystal method showed it to be a single crystal with a face-centred cubic structure. The axis of the specimen made angles of 114°·3, 80°·7, and 25°·6, with the three co-ordinate axes.

Thermal Conductivity.

The method of experiment was one in which a longitudinal heat flow was set up in a cylindrical bar and heat loss from the sides was prevented by the use of a guard ring, which in this case took the form of a coaxial shield along which the same gradient of temperature was maintained. This device greatly simplifies the experiment, for all that is required to obtain the conductivity is the rate of heat flow and the temperatures at two points in the bar a known distance apart.

The study of a single crystal by this method presents several difficulties which are not encountered when studying a bar of cast or rolled material. The crystal is so sensitive to any machining or drilling that the usual method of attaching thermocouples by pegging the wires into small holes is not permissible, for the effect of drilling such holes might possibly cause the single crystal to break up.

In these experiments copper constantan thermocouples were used, 0·3 mm. in diameter, which were held in close contact with the surface of the aluminium crystal by spring clips of the form shown in fig. 2.

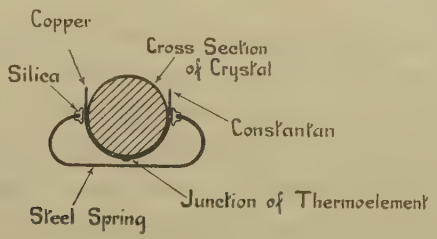


Fig. 2.—Method of attaching Thermo-elements.

* My thanks are due to Mr. W. Binks, M.Sc., of the Physics Department, for kindly making this analysis.

The amount of heat flowing to the cold end of the crystal was determined by a flow calorimeter. This was in the form of a copper spiral attached by Woods' metal to the crystal. (See fig. 3.) Water was supplied from a constant head

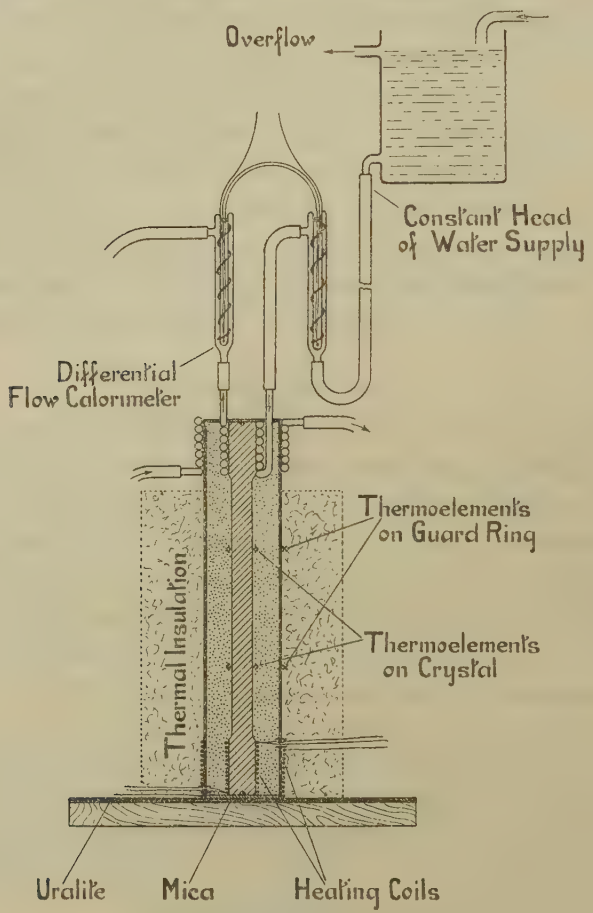


Fig. 3.

tank, and the temperature rise of the water was measured by a differential thermo-element consisting of six junctions in series made of 0·2 mm. diameter copper constantan wire. The hot end of the bar was maintained at a steady temperature by a heating coil wound directly on it and insulated with mica. The coaxial tube constituting the "guard ring" was provided with heating and cooling circuits in the same diametral planes as the corresponding ones on the aluminium crystal.

The gradient in temperature along the coaxial tube was adjusted to equality with that along the crystal by the aid of two thermocouples attached to the

surface of the tube in positions corresponding to the two thermocouples on the aluminium crystal. The space between the aluminium crystal and the "guard-ring" was filled with sil-o-cel, a very light powder of thermal conductivity 0·0002 c.g.s. units, approximately. The exterior surface of the tube was lagged with magnesia asbestos.

The apparatus employed in these tests was specially constructed, since it was necessary to use the crystal as grown from the machined bar. The cross-sectional area of the cylindrical portion of the crystal was 1·56 sq. cms. Experience has shown that it is desirable to work with bars of at least three times this cross-sectional area when making thermal conductivity measurements. It was therefore a matter of interest to determine how the absolute values of the conductivity obtained with this new apparatus compared with those given by the standard type of conductivity apparatus. For this purpose an extruded bar of aluminium (purity 99·7 per cent. Al) was taken and tested in the standard apparatus. The dimensions of the specimen were 2·54 cms. in diameter by 38 cms. in length. A piece of the same lot of material was machined down to the same dimensions as the crystal and tested in the apparatus designed for the tests on the crystal.

The results of the tests in the two forms of apparatus agreed within the limits of the possible error of experiment, being respectively $0·54_8$ and $0·55_2$ c.g.s.

For the crystal of aluminium the results given in Table I were obtained and these are shown graphically in fig. 4.

Table I.—Aluminium Crystal.

(Purity 99·6 per cent. Al.)

Temperatures.		Mean.	Value of K.
Hot end.	Cold end.		
° C.	° C.	° C.	
98·1	54·6	76·4	$0·55_1$
167·0	87·7	127·4	$0·55_2$
215·2	112·6	163·9	$0·55_7$
99·0*	54·5	76·8	$0·55_3$

* Repeat.

A chemical analysis of the material indicated that the crystal was 99·6 per cent. Al, 0·2 per cent. Si, 0·2 per cent. Fe.

It will be observed that the thermal conductivity of the single crystal is practically identical with that of the extruded bar of aluminium. In this

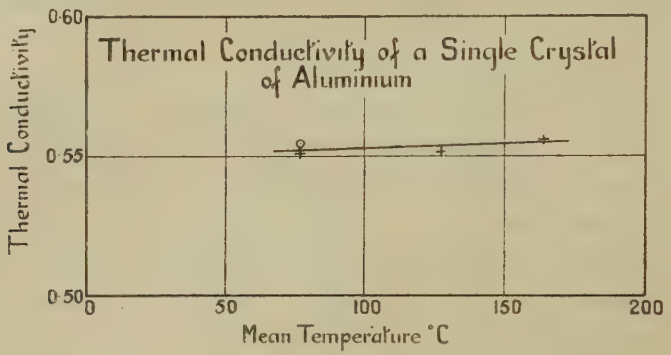


Fig. 4.

connexion it is interesting to note that Kaye and Roberts found that the "mean" conductivity of a bismuth crystal agreed with the value given by Jaeger and Diesselhorst for a rod of cast bismuth.

Another noteworthy fact is the high value of the conductivity. Jaeger and Diesselhorst in 1900 gave the value of 0·48 for aluminium. Lees in 1908 the value 0·504 for 99 per cent. purity aluminium. The writer in 1917 obtained the value 0·51 for 99·5 per cent. aluminium, so the value 0·55 for 99·7 per cent. aluminium is consistent with the view that each step in the development of the metallurgy of aluminium resulting in a purer product gives a metal of higher thermal conductivity.

It is significant, but too much weight must not be given to it, that the present tests show that the temperature coefficient of thermal conductivity of the crystal has a slight positive value. Now most alloys are characterised by a relatively large positive temperature coefficient of thermal conductivity, whilst 99·9 per cent. pure copper and 99·8 per cent. pure zinc have a slight negative temperature coefficient.

Electrical Conductivity.

The electrical conductivity of the crystal was determined by the following procedure :—

The crystal was laid on two brass knife edges which were connected to a potentiometer. Two caps were made to fit loosely over the ends of the crystal and the surrounding space filled with fusible metal. Heavy flexible leads were attached to the caps by means of which the crystal was connected in series with

a standard resistance of 0·001 ohm and a battery. When currents of 10 and 20 amperes respectively were passed through the crystal and resistance the potential drop across each was measured on the potentiometer. The observations were repeated with the current and potentiometer reversed.

The value obtained for the specific resistivity of the crystal was $2{\cdot}89 \times 10^{-6}$ ohms per cm. cube at 18° C.

Lorenz's Constant.

Many investigators have determined the Lorenz constant $K/\lambda T$, where

K is the thermal conductivity,
λ ,, electrical conductivity,
T ,, absolute temperature,

for cast or rolled pure metals, and found that it is approximately constant; the difficulties attendant on the measurement of the thermal conductivity probably accounts for some of the divergencies between the values obtained by various investigators.

For this crystal the value is $5{\cdot}4_6 \times 10^{-9}$, which falls within the range of published values 5·3 to 6·0.

The author desires to record his thanks to Mr. A. R. Challoner, Observer in the Physics Department, for his skill both in connection with the construction of the apparatus and with the observations.

On Certain Average Characteristics of World Wide Magnetic Disturbance.

By S. Chapman, M.A., D.Sc., F.R.S.

(Received May 19, 1927.)

1. This paper forms a sequel to one entitled "An Outline of a Theory of Magnetic Storms," published several years ago.* In that paper I determined the average *additional* variations in the three components of the earth's magnetic field (for observatories in magnetic latitudes† up to about 60° N.) which during times of considerable magnetic disturbance—commonly called magnetic storms —are superposed on the normal variations. The storms dealt with were such as had a commencement sufficiently definite for its epoch to be estimated to within an hour.

The average additional variations of the field were shown to be separable into two parts, one depending on "storm-time" (that is, time reckoned from the commencement of the storm), and the other being a "diurnal" variation depending on local time. Besides these average variations there are, of course, less regular features peculiar to each individual storm.

The local-time or diurnal additional variations were shown to be quite different from the diurnal variations observed on quiet days. The average diurnal variations during days of magnetic storm consist of the sum of these two diurnal variations.

2. The first object of this paper is to indicate the differences between the magnetic variations on magnetically ordinary and quiet days respectively. The magnetograph curves are less smooth and regular on ordinary than on specially quiet days, and the irregularities thus occurring on ordinary days constitute a moderate degree of magnetic "disturbance," much inferior to that which qualifies a day to be called one of magnetic storm. It is important to ascertain whether this slight magnetic disturbance differs from intense disturbance not only in degree but also in type. Mere inspection of magnetograph curves does not suffice to decide this question, which must be examined statistically, as in the determination of the average features of magnetic storms.

* 'Roy. Soc. Proc.,' A, vol. 95, p. 61 (1918). An interesting paper partly bearing on the same question has recently been published by Dr. G. Angenheister, 'Göttingen Nachrichten, Math-Phys. Kl.' (1924).

† Throughout this paper the latitudes considered will be magnetic latitudes (measured from the pole of the axis of the earth's magnetization) unless geographical latitude is expressly indicated.

It may be said at once that the evidence, over the region (up to 60° magnetic latitude) for which my former paper affords material for comparison, indicates close similarity in average type between minor and intense magnetic disturbance.

Disturbance in Middle and Lower Latitudes.

3. The simplest basis of comparison is provided by the local-time diurnal variations, illustrated in figs. 1 to 3. These figures refer respectively to the variations in the three components of magnetic force, horizontal (along the magnetic meridian), vertical (downwards), and horizontal (at right angles to the magnetic meridian, as given by the variations in west declination, measured in force units). In each case the curves illustrate the mean diurnal variation (determined in three different ways, as explained below) for the mean of one or more whole years ; the curves for the separate seasons would show differences from these annual mean curves, but such differences will not be considered in this paper. The scale of time and force is the same in all the three figures, and is indicated upon them : the force is reckoned in terms of 1γ as unit, 1γ being 10^{-5} e.m. unit of magnetic force.

In each figure there are three vertical sections, in which the individual curves are marked respectively *a*, *b* or *c*. The curves *a* show the whole average diurnal variation (for the magnetic element concerned) on quiet days, that is, on the five quiet days per month selected internationally. The curves *b* show the diurnal variations, *additional* to those in section *a*, which are observed on days of magnetic storm ; these curves are taken from figs. 3 to 5 of my former paper, and refer to a group of 40 moderate magnetic storms. The curves *c* show the diurnal variations, likewise *additional* to those in section *a*, observed on ordinary or average (*i.e.*, the mean of all) days ; these curves thus represent the hourly differences between the average diurnal variations on all days and on quiet days.

In each section *a*, *b*, *c* of each figure, five curves are shown ; these refer to different magnetic latitudes (the northernmost being at the top), as follows :—

(1) Sitka, magnetic latitude 61° ;
(2) Pavlovsk, magnetic latitude 58° ;
(3) Pola, Potsdam, Greenwich—mean magnetic latitude 51° ;
(4) Zikawei, San Fernando, Cheltenham, Baldwin—40° ;
(5) Batavia, Porto Rico, Honolulu—22°.

The curves (3)–(5) in section *c* have been determined respectively from Greenwich, Cheltenham, and Honolulu only, in order to save labour.

The curves b and c represent the average diurnal effect of (b) moderately strong and (c) relatively weak magnetic disturbance respectively, superposed

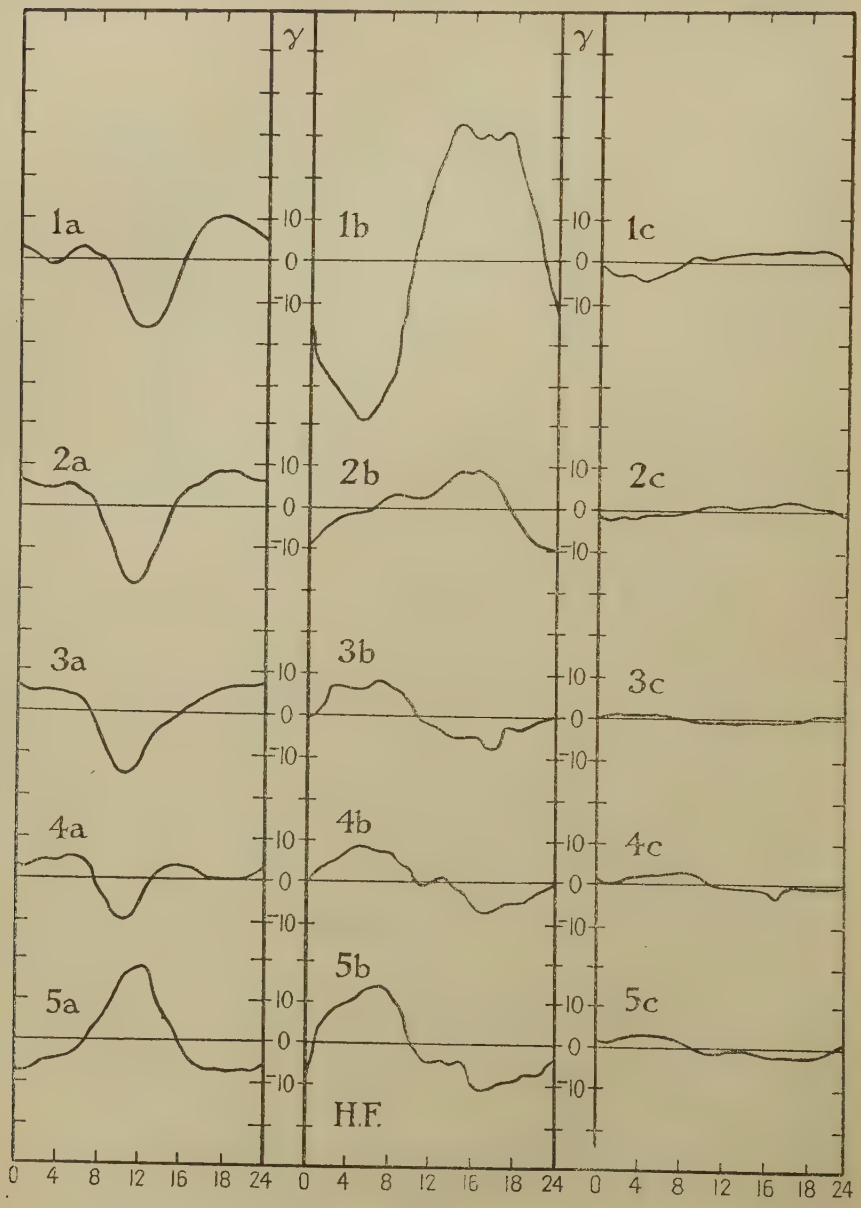


Fig. 1.

on the permanent diurnal variation a, which on quiet days appears alone. In spite of various minor irregularities, many of which would disappear if averages

over a larger amount of material had been taken, the 15 independent curves *b* show a considerable degree of similarity with the corresponding 15 curves *c*,

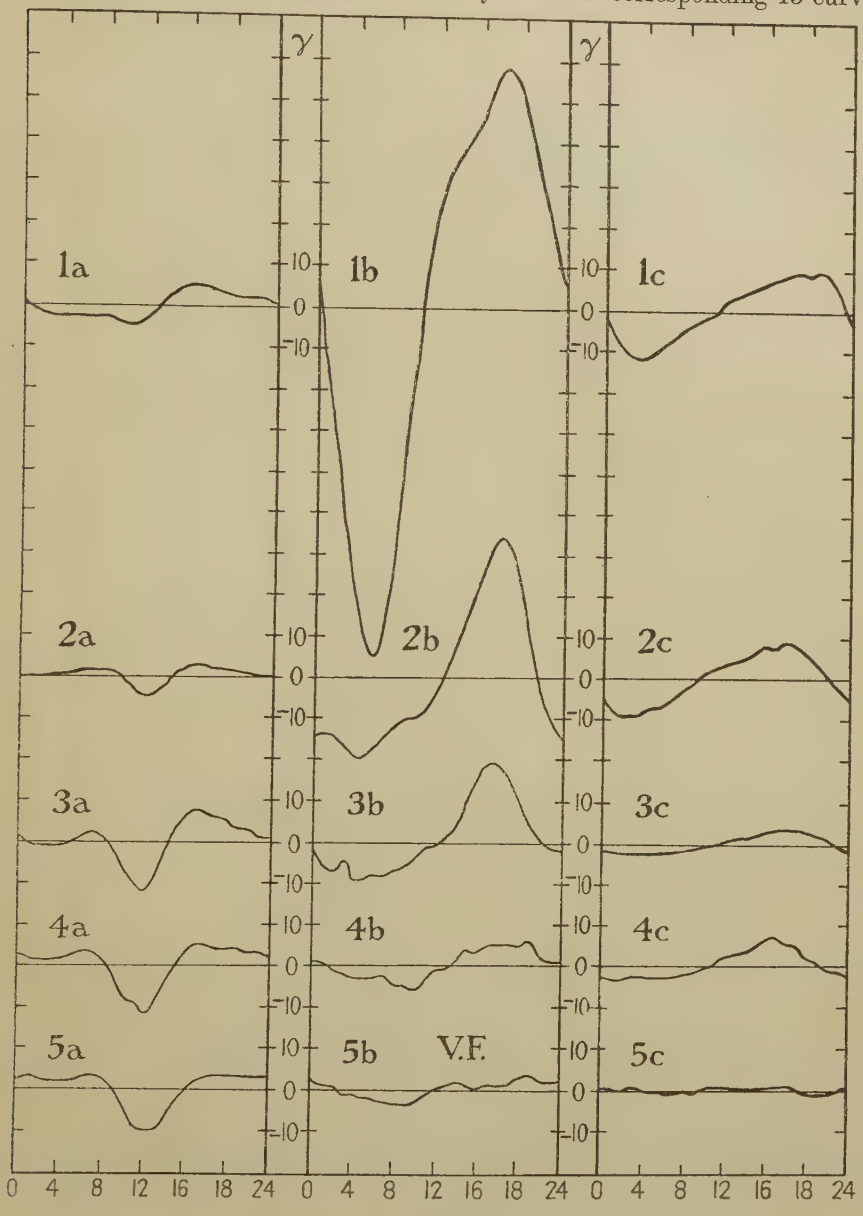


FIG. 2.

while the curves *b* and *c* are, obviously, quite different from the curves *a*.* For example, in fig. 1 there is a reversal of type in section *a* in a low latitude

* The curve 1*a* in fig. 2 is the only exception : this is referred to later in § 9.

(about 30°, between curves 4 and 5), while in sections *b* and *c*, though there is also a reversal of type, it occurs in a much higher latitude (about 55°,

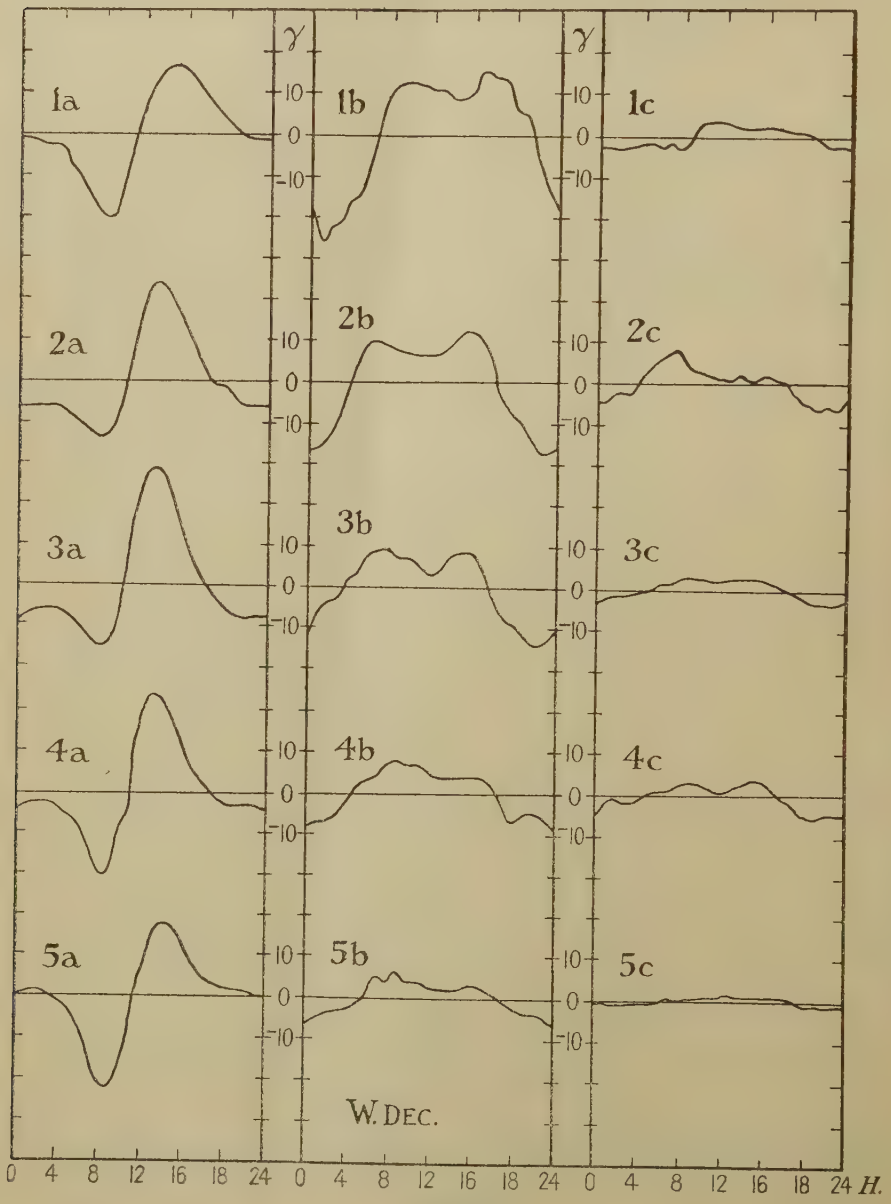


FIG. 3.

between curves 2 and 3); moreover, in section *a* the extreme departures from the mean occur in the middle of the day, at about the time when curves *b*

and *c* pass through the zero line. The latter contrast is shown also by fig. 2, though the two figures differ greatly in other respects. In fig. 3 the curves *a* pass through zero at about midday, while the curves *b* and *c* pass through zero at about 6 h. and 18 h.

These three figures summarize a considerable mass of evidence, and one that could be easily extended, as, for example, by taking more stations, by separately considering years of sunspot maximum and sunspot minimum, and by considering the seasons separately. Some of these extensions have, in fact, been made, but the detailed results need not be given here. The evidence, as in figs. 1 to 3, points definitely to the conclusion that the average additional magnetic variations during magnetic disturbance, which are superposed on the quiet-day variations, do not differ much in *type*, as the intensity varies over a wide range.

4. The above evidence relates to one part only of the additional magnetic field present during disturbance. The other part is that which depends on storm-time, and represents the effect of the disturbance, at successive times after the commencement, averaged round each parallel of latitude. The storm-time part is specially strong and definite in the horizontal force (*cf.* fig. 1 of my former paper). If the conclusion stated in § 3 is correct, similar but smaller changes should accompany minor disturbance. Unfortunately this cannot be tested in the same detailed way as for the local-time variations considered in § 3. The difficulty lies in the determination of storm-time, reckoned from the beginning of the disturbance. This is often easy in the case of magnetic storms, because they are well-marked individual phenomena ; but it is usually hard to distinguish whether a prolonged period of minor disturbance is made up of one or many individual disturbances.

It is possible, however, to apply a partial test to the above conclusion, having reference to the storm-time changes. The local-time changes considered in § 3 are measured from the daily mean value of the magnetic element as origin, and are independent of the absolute value of this mean. The storm-time changes, on the other hand, directly affect the daily mean value, and the way in which they will do so on ordinary days, if they are then of the same type as in magnetic storms, can be inferred on the reasonable assumption that, when a large number of more or less disturbed days are considered, each hour of the day will experience approximately the same average storm-time effect. For example, during the first few hours of a magnetic storm the horizontal force is generally increased (taking the average round a parallel of latitude), but subsequently it experiences a much larger and more prolonged decrease, which dies

away slowly : thus, the average storm-time effect on the horizontal force should be a decrease in the daily mean during disturbance. The magnitude of the decrease should have a maximum at the equator (*cf.* fig. 1 of my former paper), and should diminish with increasing latitude to about half the equatorial value, at about 50° latitude.

The vertical force storm-time changes are much smaller, and opposite in sign, but otherwise follow a similar course. The averaged effect on the daily mean value of the downward vertical force should be a very slight increase, in moderate northern latitudes : the effect should vanish at the equator and be reversed (like the whole vertical force) in the southern hemisphere.

The west declination shows no appreciable storm-time changes in lower and middle latitudes, and therefore the daily mean value of the declination should not be affected by disturbance.

These expectations, based on the conclusion of § 3 that the average characteristics of slight and intense disturbance are similar in type, are well borne out by observation, as the following table shows. It gives the differences obtained by subtracting the daily mean values of the three magnetic elements, for *quiet* days, from the corresponding daily means for *all* days. The stations are those already considered in § 3 (their magnetic latitudes are indicated in brackets) : the results are expressed in terms of the unit 1γ.

	H.F.	V.F.	W.Dec.
(1) Sitka (61°)	— 3	— 3	0
(2) Pavlovsk (58°)	— 5	0	0
(3) Greenwich (55°)	— 4	0	0
(4) Cheltenham (49°)	— 7	1	0
(5) Honolulu (21°)	—10	1	1

Though these results do not all refer to the same series of years, a different choice would not alter the general indications of this table. The changes observed are small except in the case of the horizontal force, where they are systematic, and in good agreement with expectation. The results are given to the nearest unit, and are probably accurate to within about 1γ.

In addition to this evidence favouring the view that also as regards the storm-time changes slight and intense disturbance are similar in type, there is the direct evidence afforded in fig. 6 of my former paper showing that the type of the storm-time changes during intense disturbance is but little altered over a range of intensity varying in the ratio 1 to 4. There appears to be a

quicker progression through the storm-time changes, the more intense the disturbance, but the similarity of type is the outstanding feature.

5. From §§ 3, 4 it appears that, in addition to the main magnetic field of the earth, and the varying field which is manifested by the diurnal magnetic variations on quiet days, there exists, from time to time, a further "disturbing" magnetic field, which has certain definite average characteristics in regard to its space distribution and temporal changes: and that, over a wide range of variation of intensity of this field, these average characteristics do not vary greatly in *type*.

This conclusion relates to world-wide magnetic disturbance, and not to isolated local or brief disturbances (such as "bays"), except in so far as these are associated with and indicative of the general magnetic disturbance which distinguishes ordinary from quiet days.

Disturbance in high latitudes.

6. The evidence so far described relates to latitudes up to about 60°: the data for polar regions will next be considered. The observational material is more limited than for lower latitudes, particularly in respect of the length of record at any individual station: the period during which continuous magnetographic registration has been carried on at polar stations has rarely exceeded two years. Consequently the polar observations for days which in lower latitudes are counted as highly disturbed are too few to enable average results for such days to be obtained for polar latitudes, corresponding to those of my former paper relating to lower latitudes.

In particular, it does not seem possible at present to determine the course of the storm-time changes in the polar regions. But it is possible to compare days of different degrees of slight disturbance (reckoned by polar standards, which, of course, are much higher than in lower latitudes), as was done in §§ 3, 4 for non-polar stations, so that the detailed local-time variations, and the averaged storm-time variations, can be found for polar latitudes. The results so obtained will, of course, have reference to the same relatively low degree of disturbance as in the case of the curves *c* in figs. 1 to 3.

7. The averaged effect of the storm-time changes will first be considered, as shown by the difference between the daily mean values of the magnetic elements on all days and on quiet days. The data obtained at several arctic stations during the international expeditions of 1882–3 are available for this purpose, and also the results of recent Antarctic expeditions, and the hourly values for 1914, 1915 recorded at the new permanent observatory at Sodankyla, in

Finland, a little to the south of the auroral zone. Where the published observations include results for quiet days separate from those for all days, these two sets of results have been used : where quiet-day results are not quoted, they have been computed for sets of quiet days chosen by the criterion of relatively small daily range. The results so obtained may not be strictly comparable *inter se*, but are sufficient for the present purpose.

The difference between the all and quiet day means for each station is a vector with three components : in polar regions the horizontal component of this vector does not in general lie so nearly along the direction of the whole horizontal force as it does in lower latitudes, and, moreover, the geographical direction of the horizontal force varies considerably. It is, therefore, desirable to indicate graphically the direction of the horizontal disturbance vector, as in fig. 4.

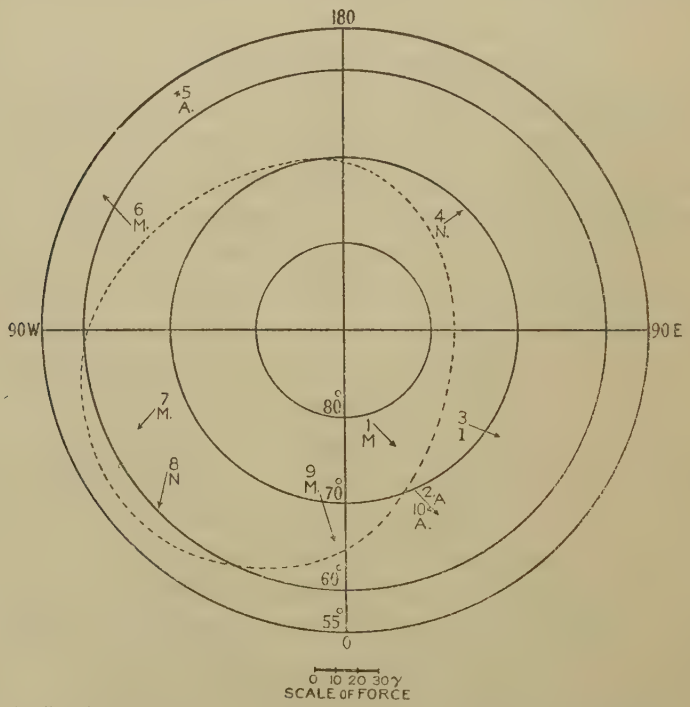


Fig. 4.—1. Cap Thordsen. 2. Bossekop. 3. Nova Zembla. 4. Ssagastyr. 5. Sitka. 6. Fort Rae. 7. Kingua Fiord. 8. Godthaab. 9. Jan Mayen. 10. Sodankyla.

M = V.F., Morning Maximum. I = V.F., Irregular.
A = V.F. Afternoon Maximum. N = V.F. No Record, or Unreliable.

In this figure the situation in latitude and longitude is represented for ten observatories by a point in each case, and from each point a line (having an arrow-head at the further end) is drawn indicating in magnitude and

direction the difference between the mean horizontal component of magnetic force on all and on quiet days. The zone of maximum auroral frequency, as drawn by Fritz, is also shown (dotted). It appears that, on the whole, these horizontal disturbance vectors diverge from a point or small region near the centre of the auroral zone, which is also approximately the pole of the axis of magnetization of the earth ; their distribution is by no means symmetrical about the geographical pole.

The change in the mean horizontal force thus consists approximately of a reduction in the component of the force along the meridians through the magnetic axis, and in this respect it resembles the corresponding horizontal-force change in lower latitudes (§ 4) ; but the magnitude of the reduction is larger in polar regions, ranging up to about 20γ, whereas in lower latitudes the maximum, which occurs at the equator, is about 10γ, though these figures may not be exactly comparable owing to the different years and days from which they are drawn.

The average storm-time change of force in the horizontal plane is thus everywhere a decrease, which has a maximum at the equator, whence its value decreases towards a minimum at about 60° latitude, afterwards increasing rapidly towards the auroral zone. The data do not suffice to show clearly how it varies within the zone, but it may be expected to decrease again, to zero, at the centre of the zone.

The corresponding changes in the vertical force are more difficult to ascertain, since not all the polar stations of 1882–3 were equipped with reliable (or, in some cases, with any) vertical force magnetographs. But examination of the available data makes it appear that at Fort Rae, Nova Zembla, and Kingua Fjord the all-day mean is distinctly higher than the quiet-day mean, by about 18γ, 19γ, and 6γ respectively. At Bossekop, Sodankyla, and Sitka the change is negative and small, viz., — 4, — 4, — 3γ. At Cap Thordsen and Jan Mayen the change appears to be small and its sign cannot be reliably determined from the existing data.

The large V.F. changes, for which numerical values have just been quoted, are probably reliable as regards sign and order of magnitude, at least ; they much exceed the corresponding V.F. changes in lower latitudes, which are small and positive from the equator (where the change is zero) up to about 55° northern latitude, where they change sign (to negative) and increase numerically towards the auroral zone. A further change of sign, to positive values, occurs at or within the auroral zone, where the large differences above noted are found. Further within the zone it seems likely that the differences decrease again

towards the axis of magnetization, but data on this point are not available. These vertical force changes are, therefore, more complicated than those in the horizontal plane.

In the Antarctic the results for Cape Evans (77° 38′ S., 166° 24′ E., geographical), reduced and discussed by Dr. Chree,* show that the "all-minus-quiet day" mean difference in the (upward) vertical force in an increase of 5γ; this result is confirmed by the still larger difference of 12γ, with the same sign, shown by the "disturbed-minus-all day" means (from the five most disturbed days per month). Thus within the southern auroral zone (which encircles Cape Evans) disturbance increases the numerical value of the vertical force, just as in the case of the northern polar region.

8. The local-time (diurnal) additional variations will next be considered. These have been discussed by Lüdeling† and van Bemmelen,‡ and the results here to be quoted, though derived from an independent examination of the data, are in substantial agreement with theirs.

Fig. 2 (curves *b*, *c*) shows that the additional diurnal variation in the vertical force preserves a constant phase from the equator to as far north as Sitka, with a morning minimum and an afternoon maximum; the amplitude increases greatly with latitude. The increase persists beyond the (magnetic) latitude of Sitka—where the range, for all-minus-quiet days, is about 20γ (*cf.*, fig. 2, curve 1*c*)—to Bossekop, where the range exceeds 100γ; the curves for Bossekop and other polar stations are illustrated in fig. 5. These curves (except in the case of that for Nova Zembla) are of similar very large range, but the most remarkable feature is that their phases are all opposite to that for Bossekop and the stations of lower latitude; the curve for Nova Zembla is transitional between the two series. The reversal of phase would seem to occur within a narrow belt of magnetic latitude adjacent to the auroral zone. Such a reversal, in a region where on the two sides of the dividing line the range of the diurnal variation (of V.F.) is so large, constitutes what is perhaps the most striking of all the average characteristics of the field of world-wide magnetic disturbance. (§ 5).

* C. Chree, 'Roy. Soc. Proc.,' A, vol. 104, p. 165 (1923).

† 'Terrestrial Magnetism,' vol. 4, p. 245 (1899).

‡ *Ibid.*, vol. 8, p. 153 (1903). Dr. van Bemmelen also investigated the alterations in the daily mean values of the magnetic elements in the days succeeding magnetic storms, and showed that the azimuth of the disturbing force in the horizontal plane was fairly constant at each station. For the arctic region the azimuths of the horizontal disturbing vectors in fig. 4 of the present paper agree closely with those found by Dr. Bemmelen for his different but related vectors: *cf.* "Die Erdmagnetische Nachstörung," 'Met. Z.,' 1895, p. 321, and also 'Terrestrial Magnetism,' vol. 5, p. 123 (1900).

The additional "all-minus-quiet day" diurnal variations in the horizontal plane are likewise of considerable magnitude in polar regions, the range being of the order 50γ, as for the vertical force. This is far greater than in lower latitudes, and even than at Sitka, in magnetic latitude 61° (fig. 1, 1*c*). In passing northwards from Sitka and crossing to within the auroral zone the diurnal variation of the force in the horizontal plane experiences a striking change of type, which is not simply a reversal, as in the case of the V.F. The change is best shown by the vector-diagrams of the horizontal-force additional variation. At Sitka the diagram still bears some resemblance to the roughly oval form, elongated in the direction transverse to the magnetic meridian, shown at Greenwich and other stations in similar latitudes. But for stations quite near to the auroral zone, like Ssagastyr, Sodankyla, Nova Zembla, and Bossekop, in magnetic latitudes 61°, 66°, 64°, 67°, the diagram is very narrow in this direction (*i.e.*, along the zone), and elongated in the direction normal to the zone; *cf.* fig. 6. The maximum poleward force occurs at about 16 h. or 18 h., and the opposite minimum at about 2 h. On passing well inside the zone the vector diagram again becomes oval, indeed nearly circular; this is illustrated by the diagrams for Kingua Fjord and Cape Evans, in magnetic latitudes 79° N. and 83° S.; note that these two curves are described in opposite senses (*cf.* § 17). The curves for other stations in or near the zone, like Cap Thordsen, Fort Rae, and Point Barrow, are of an intermediate and less simple type.

The horizontal vector force diagrams, for all-minus-quiet days, are given to scale in fig. 6 for Sitka, Sodankyla, Bossekop, Kingua Fjord and Cape Evans. In each case the magnitude and direction of the horizontal force disturbance vector illustrated in fig. 4 is indicated, by a line drawn from the origin of the diagram. This enables the orientation of the diagram, and therefore of the diurnal disturbance vector at any local time, relative to the auroral zone, to be inferred by comparison with fig. 4. It also shows the relative magnitude of the disturbance change in the mean value of the force in the horizontal plane, and of the diurnally varying departure from that mean. It should be carefully noted that the diagrams in fig. 6 are drawn to different scales, in order to render them of similar size and so facilitate comparison between their different forms.

9. The average characteristics of disturbance in polar regions, considered in §§ 6–8, are those associated with relatively slight disturbance, but the results described in §§ 2–4 suggest that the distribution and development of the additional disturbance field in polar regions, as well as in lower latitudes, may remain fairly constant and independent of the intensity of the disturbance.

There is, indeed, some polar evidence for this conclusion; this evidence will be briefly reviewed.

It has been seen that the additional disturbance diurnal variations in polar regions are very intense (the range being of the order of 50γ), as compared with the corresponding variations in lower latitudes. Even on relatively quiet days, such as the five international quiet days per month, there remains a certain amount of disturbance, with its associated disturbance diurnal variation. In middle and low latitudes this is small compared with the normal quiet-day variation, but it is still considerable in high latitudes, sufficiently so to be able to mask partly or wholly the residual diurnal variation corresponding to an ideally quiet day. Curves 1 *a*, *c* of fig. 2, illustrate this: Curve 1 *c* is of larger amplitude than 1*a*, and the latter appears to be compounded of two variations, one being a residue of disturbance diurnal variation, similar to but smaller than 1*c*, and the other a reduced version of the normal quiet-day vertical force variation shown in fig. 2, curves 3, 4, 5*a*.

Within the auroral zone the normal quiet-day variation probably becomes insignificant compared with the disturbance diurnal variation, even in the reduced form of the latter on quiet days. If this is so, the similarity of type between the whole diurnal variations, whether on quiet, average, or more than usually disturbed days (5 per month), which has been clearly demonstrated by Dr. Chree in reference to Antarctic magnetic records, constitutes a partial confirmation of the view that the general character of the disturbance field, in polar as well as in lower latitudes, does not vary greatly over a wide range of intensity.

It may be added that this view is consistent with the possibility (for which there appears to be some observational evidence) that at an individual polar station near the auroral zone there may be considerable change of type in the disturbance diurnal variation, or in the sign and magnitude of the disturbance change in the mean magnetic force, as the intensity of disturbance increases. This is because of the complicated structure of the average disturbance field in polar regions, and its close relationship with the auroral zone. This zone appears to broaden and move towards the equator during periods of intense disturbance, so that a station which ordinarily is on the equatorial side of the zone may be during magnetic storms under, or on the polar side of, the zone. If so, the disturbance changes in the magnetic field at the station during slight disturbance may be radically different from those during magnetic storms, even though, in relation to the zone, the character of the disturbance field may be similar in the two cases.

The station Nova Zembla seems to exemplify this possibility. The average vertical force variation for this observatory (fig. 5) is of transitional type, suggesting that the observatory is nearly under the dividing line between the two types of variation shown in fig. 5 (this dividing line is probably identical with the auroral zone). But on many individual days the V.F. variation is of definite type, sometimes with morning maximum and afternoon minimum, and at other times with the reverse phase. Mr. T. T. Whitehead, who has kindly assisted me in other ways in the preparation of this paper, examined the days of these two classes at Nova Zembla. Out of about a year's observations, 100 days of morning V.F. maximum, and about 112 days of afternoon maximum, were found. On the remaining days either the record was faulty, or the maximum and minimum V.F. both fell in the forenoon or afternoon with less than six hours' interval; most of the very highly disturbed days were in this class. The mean range of the horizontal force variation on the above two classes of days was determined, and was found to be distinctly greater on the days of morning than of afternoon V.F. maximum (429γ as against 268γ). This suggests that on the more disturbed days Nova Zembla may be inside the auroral zone, while on quieter days it is outside it. It would be of interest to examine the records of other observatories, near the auroral zone, in a similar way.

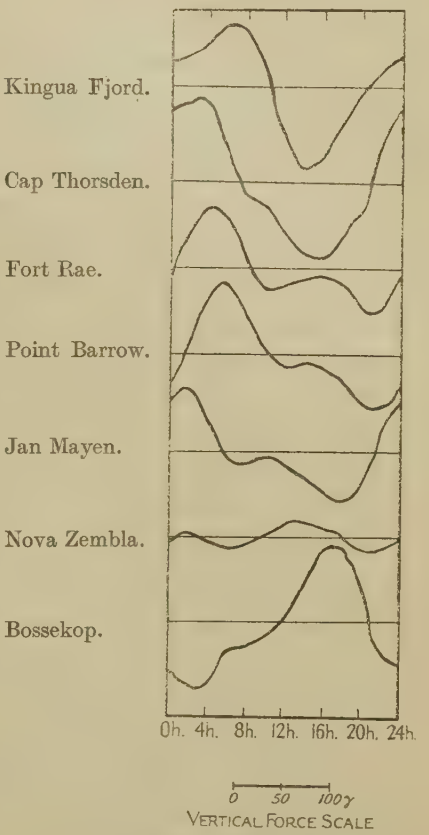


FIG. 5.—Annual Mean Diurnal Variation of Vertical Force in Polar Regions.

The Electric Current System.

10. The preceding sections summarize what appear to be the principal average characteristics of the field of world-wide magnetic disturbance. As in my former paper, the attempt will now be made to infer the immediate cause of this field. The most natural supposition is that the field is due to a system

of electric currents, part being situated in the upper atmosphere, and part within the earth, the latter being secondary currents induced by the variations of the former. This view is supported by the known facts about auroræ, and by analogy with the theory, incomplete yet already to a large extent successful, of the ordinary solar and lunar diurnal magnetic variations. It has also been confirmed, in as direct a way as is ever likely to be possible, by Birkeland's study* of large perturbations of the horizontal and vertical components of magnetic force at adjacent stations in the Arctic region. He found a number of cases in which there was clear evidence of a linear current flowing overhead between two stations, the horizontal force perturbations being of the same sign, and the vertical force perturbations of opposite sign ; the estimated heights of the centre-lines of such currents varied from 150 km. upwards.

In my former paper diagrams were given of current-systems that would account in a general way for the data there described, which were confined to phenomena in middle and lower latitudes, up to about 60° (that is, to the range of facts considered in §§ 2–4 of the present paper). In the main those diagrams are still valid for this region, but they require important modifications in the polar regions, where, as has been shown, the disturbance field is not only more intense, but also more complex in structure, than elsewhere.

The complications of the disturbance field in high latitudes may be divided into two classes, one set being associated with the very existence of a band or zone (the auroral zone), at ordinary times comparatively narrow, within which the atmosphere receives from the sun the electric particles that are the primary cause of magnetic disturbance. The corresponding complications of the disturbance field, whatever their nature, may be called the symmetric-zonal features of the field: they would exist even if the earth's magnetic axis coincided with its geographical axis, when, as the Birkeland-Störmer theory of auroræ indicates, the auroral zone would be symmetrical about the earth's axis of rotation, and auroral phenomena would, on the average, follow the same course at corresponding local times all round any circle of geographical latitude. The fact that the magnetic and geographical axes do not agree, so that the auroral zone is not centred at the north pole, introduces a second set of complications, which may be called the asymmetric features of the field. The angular distance between the north pole and the pole of the magnetic axis (or the centre of the auroral zone) is comparable with the angular radius of the auroral zone, and

* Birkeland, 'The Norwegian Aurora Polaris Expedition,' 1902–3, vol. i, § 76, p. 306. Currents of the order of 10^6 amperes were indicated during magnetic storms not of outstanding intensity.

the divergence must produce important consequences in the disturbance field in these regions where the structure of the field is so highly differentiated ; in low latitudes, on the other hand, the asymmetric features of the disturbance field are relatively unimportant. One chief complication in polar regions is that the disturbance diurnal variation will not be the same at similar local times at different stations along a circle of either geographical or magnetic latitude.

The zonal and the asymmetric features of the polar disturbance field cannot be separated with certainty on the basis of the few brief series of polar observation so far available, and it is desirable to obtain more data by means of a closer network of permanent or temporary polar observatories, from which the detailed structure and variations of the disturbance field may be determined. But the zonal features of the field seem on the whole to be more important than the asymmetric ones, and in this section they alone will be considered. Current systems will be proposed which relate to an ideal earth having coincident magnetic and geographical axes ; these currents would give rise to a magnetic field that possesses the most important features of the actual disturbance field, when comparison is made between a station on the ideal earth, in a given situation relative to the ideal auroral zone symmetric about the axis, and a station on the actual earth in a similar situation relative to the actual auroral zone. The hypothetical and the actual disturbance fields will, of course, not agree in all particulars, but the ideal current system, and its hypothetical field, will be of value according to the extent to which they reproduce the observed features of the disturbance field ; and a study of the differences between the hypothetical and observed variations, when sufficient data become available, should then assist in the determination and explanation of the asymmetric features of the field.

11. In the foregoing discussion, the disturbance magnetic field has been analyzed into two parts, the diurnal or local-time part, and the storm-time part. The current system responsible for the whole field can be similarly analyzed, since the magnetic fields of superposed current systems are superposable.

The ideal current system proposed as corresponding to the symmetric or zonal features of the *diurnal* part of the disturbance field is indicated in fig. 7. This shows an elevation of the atmospheric current systems as shown from the sun, and also a plan as seen by a distant observer looking along the earth's axis towards the North Pole. The ideal auroral zones, or zones of great current intensity nearly coinciding with them, are indicated by $\alpha\beta$, $\alpha'\beta'$, the former being shown as a circle in plan. The currents along and in the polar

caps within these zones are supposed to be much more intense than those between the zones in middle and low latitudes, so as to account for the greater intensity of the disturbance field in polar regions. Most of the currents flowing along the zone (in opposite directions in the two halves of each zone) are supposed to complete their paths by flowing across the polar caps within the zone, but part of the zonal currents may also complete their circuits over the middle belt of the earth. The distribution of the current system in space is supposed to be constant relative to the reference frame determined by the earth's axis and the radius vector from the earth's centre to the sun ; the earth revolves relative to this frame, and relative to this current system ; in the atmosphere above any particular locality on the earth the currents are continuously modified, to the form characteristic of the situation of the locality relative to this reference frame ; this situation is determined, of course, by

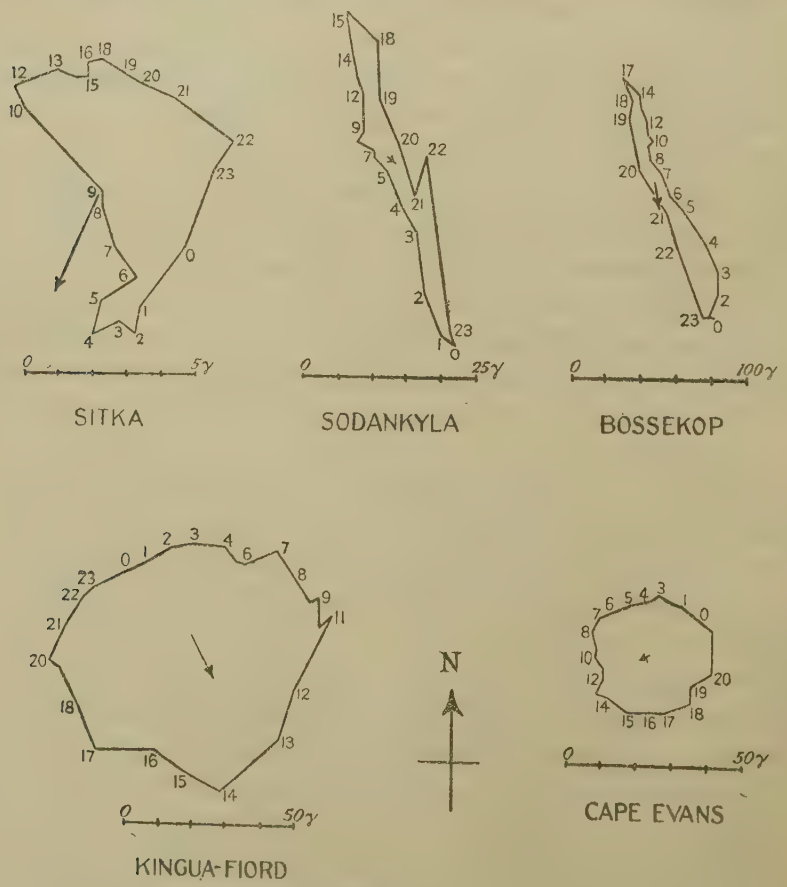


Fig. 6.

the latitude and local time of the station. The magnetic field of the current system is observed as a field which varies diurnally at each station, according to local time.

12. The nature of these diurnal variations in the three components of magnetic force at a point at ground level—that is, below the current system—can easily be inferred from fig. 7. At the equator, and on both sides up to the latitude (N. or S.) at which the current circuits between the two zones shrink to points, the horizontal north force will have its maximum at 6 h. and its minimum at 18 h. local time; above these latitudes, up to the zones $\alpha\beta$, $\alpha'\beta'$, the north force variation will be reversed. This agrees with the general character of the curves *b*, *c* in fig. 1. The variation of west force or declination will be of opposite phase on the two sides of the equator, without further reversal up to the zones $\alpha\beta$, $\alpha'\beta'$; the west force in northern latitudes, up to $\alpha\beta$, will have its maximum at 12 h. and its minimum at midnight. These inferences from fig. 7 are in general agreement with the curves *b*, *c* of fig. 3. The vertical force diurnal variation will likewise be of constant phase from the equator up to each zone, and of opposite sign on the two sides of the equator, but the maximum downward force in northern latitudes up to $\alpha\beta$ will occur at 18 h., and the minimum at 6 h., while the daily range will increase from zero at the equator to a maximum at a latitude only slightly less than that of the centre line of the zone $\alpha\beta$. These inferences are in general agreement with the observed facts summarized in the curves *b*, *c* of fig. 2.

The above is the range of facts on which the corresponding current-system diagram of my former paper (fig. 7) was based; that diagram accounts fairly well for these facts except as regards the rapid increase of the range of the diurnal variation of vertical force above 50° latitude (in particular, at Sitka). It does not agree at all with the data for the polar regions (§§ 6–9 of this paper), since it took no account of the intense current zones $\alpha\beta$, $\alpha'\beta'$, included in the present fig. 7.

13. The form of the diurnal variations in polar regions, corresponding to fig. 7, will next be considered. The phase of the diurnal variation of vertical force will remain constant up to the centre-line of the zone $\alpha\beta$, beyond which the phase will be reversed; also the range will have a maximum at a latitude slightly below $\alpha\beta$, and will decrease rapidly from that latitude to zero at stations directly under the zone. Further north still, the range of the reversed variation will rapidly rise to another maximum a little north of the zone, and from there will gradually decrease to zero at the pole or centre of the zone.

These inferences are in general agreement with the facts summarized in

fig. 5, since an oval curve can be drawn on the map of the Arctic region so as to separate the stations having vertical force variations of opposite phase, in such a way that those stations within the curve have maximum downward vertical force in the forenoon, while at those outside the curve it occurs in the afternoon. The curve must be drawn (*cf.* fig. 4) so as to pass to the north of Fort Rae and Bossekop, and nearly over Nova Zembla. It will therefore not agree exactly with the zone of maximum auroral frequency as drawn by Fritz, though its departure from the latter is not great. The magnetic data are unfortunately insufficient to indicate the complete course of the curve referred to; for this purpose it is desirable that at various points along the auroral zone there should be pairs of magnetic observatories in nearly the same magnetic meridians, and only a few degrees of latitude apart, on opposite sides of the zone. In view of the special interest attaching to the vertical force variations in polar regions, it is unfortunate that observations of this element are particularly scanty in these latitudes. The future outlook in this respect is now somewhat more hopeful, and for the present it is at least possible to say that the existing data seem to be in good accord with the indications of fig. 7.

14.1. As regards the polar diurnal variations of magnetic force in the horizontal plane, the observed variations of the two components, north and west, have been accounted for in § 12 up to the latitude of Sitka. As the current-zones are approached, the horizontal force variation will become more and more directly transverse to the zone, the west-component variation decreasing in range till it vanishes under the zone, the north-component variation remaining the same in phase (morning minimum and afternoon maximum) and increasing greatly in magnitude. The vector diagram of the horizontal force variation will thus become rectilinear, with its direction normal to the zone, and with its north and south elongations at 18 h. and 6 h. respectively.

14.2. On crossing the zone, the north-component variation will preserve its phase, but its range will begin to decrease; the west-component variation will reappear with reversed phase. The horizontal-force vector diagram will therefore again become oval, but will be described in the counter-clockwise direction, instead of in the clockwise direction (as is the case just south of the zone $\alpha\,\beta$); the horizontal force-vector at stations just within the zone will be southerly at 6 h., easterly at 12 h., and so on.

14.3. Still further within the zone the north-component variation will vanish and reappear with reversed phase; the west-component variation will have the same phase everywhere within the zone. The direction of description

of the vector diagram will thus be again reversed—the third reversal in all in proceeding northwards from the equator (the three occurring at about the latitudes 55°, 65°, and 75°, the middle latitude being that of the zone $\alpha\beta$ itself): the direction of description in the central region within the zone will therefore be clockwise, the force-vector being northerly at 6 h., easterly at 12 h., and so on. A simple way of regarding this variation, in the vicinity of the pole, is to consider it as due to the rotation of the earth within a stationary magnetic field, which near the poles is nearly uniform and horizontal: relative to the earth, therefore, the variation at any particular station near the pole will appear as the uniform rotation of a nearly constant vector, in the direction opposite to the rotation of the earth. The constant vector, according to fig. 7, will be in the meridian plane normal to the radius vector to the sun, or, otherwise stated, in the meridian plane of the 6 h. and 18 h. meridians of local time. This gives rise to north and west component variations with phases as just described; also the vector diagram will be nearly circular. The reference here is to the arctic cap; in the antarctic the horizontal component of the force near the pole will be opposite in direction to that at the south pole, as fig. 7, indicates, and the rotation of the earth will cause the vector diagram at stations near the south pole to be described in the anti-clockwise direction, the vector being northerly at 6 h., westerly at 12 h., and so on.

15. The hypothetical horizontal force diurnal changes, or vector diagrams, in the polar regions, deduced from fig. 7 and described in § 14, are in the main in fair accord with observation, though there are some discrepancies. The transition from the vector diagram at Sitka to the rectilinear vector diagram at Bossekop, which is under or very near to the auroral zone, is in accord with § 14, though the observed phases do not always agree with the hypothetical phases.* In high northern and southern magnetic latitudes, such as at Kingua Fjord and Cape Evans, the observed vector diagrams are nearly circular, are described nearly uniformly, and in the directions described in § 14, but again there are phase-differences of about 3 hours between the actual and the hypothetical diagrams. Such differences are not surprising in view of the very considerable asymmetry of the auroral zones with respect to the geographical poles. The same cause may explain the lack of observational data (so far

* In the diagrams for Bossekop and Sodankyla, for example, the extreme elongations are at 0 h. and 15 h. or 17 h. instead of at 6 h. and 18 h. as suggested in § 14.1; the diagram for Ssagastyr (not included in fig. 6) has its extreme elongations at 4 h. and 18 or 20 h. Thus at different stations near the auroral zone the maximum current along the zone seems to occur at different hours of local time: this is an example of the consequences to be expected from the asymmetry of the zone about the geographical axis.

as I am aware) confirming the predicted reversal, between the auroral zone and the pole, of the direction of description of the vector-diagram, though it seems possible that a fuller survey may reveal stations not far within the zone which have a counter-clockwise vector diagram. In any case, it may reasonably be claimed that the relatively simple current system of fig. 7 suffices to account for a large proportion of the considerable range of facts summarized in figs. 1-3, 5, 6.

16. The second part (*cf.* § 11) of the idealized current system will next be considered, that, namely, which is responsible for the storm-time part of the varying disturbance field. By virtue of its definition and mode of derivation this part is essentially symmetrical about the earth's axis ; its intensity and its distribution in latitude vary very materially during the course of a storm, as is shown by the storm-time curves, for middle and low latitudes only (up to about 60°), given in fig. 1 of my former paper. Since it does not seem possible as yet to deduce similar curves for the polar regions, it is necessary to confine our attention to the time-average of the storm-time changes, so that the corresponding current-system diagram, fig. 8 of this paper, will refer only to the time-average of the actual varying storm-time part of the disturbance field ; it will, therefore, completely fail to represent the current system of this part of the field during the early hours of a large disturbance, when the storm-time changes are of opposite sign (in middle and low latitudes, at least) to the average storm-time changes.

The averaged storm-time changes have been described in §§ 4, 7 ; they constitute a much more limited set of facts than those for which fig. 7 was required to account. The corresponding current-system, shown in fig. 8, consists of currents which are everywhere in the same direction, along circles of latitude from east to west, but concentrated, with much greater intensity than elsewhere, along the same zones, $\alpha\beta$, $\alpha'\beta'$, as in fig. 7. The associated magnetic field will everywhere be directed southwards, and will be most intense in and near the current zones ; this agrees with the diminution of north force observed everywhere in middle and low latitudes ; and with the diminution, near the auroral zone, of the component of horizontal force normal to the zones. The vertical force will clearly vanish at the equator, and be large, and of opposite signs, adjacent to and on opposite sides of the zones. Just south of $\alpha\beta$ it will be upwards, and just north it will be downwards. These indications are in general accord with the facts, though the vertical force data are too scanty, in polar regions, to test them properly. The magnitude and sign of the hypothetical vertical force variations in low latitudes will depend on the precise distribution in latitude of the current intensity in fig. 8.

Fig. 8 is similar to the storm-time part of the current system described in my former paper (§ 8 (*a*), *loc. cit.*) except as regards the distribution of current intensity. In that paper the current intensity was described as being nearly uniform in middle and low latitudes, but gradually diminishing towards high latitudes ; the existence of the zones of large current intensity, $\alpha\beta$, $\alpha'\beta'$, was not indicated, owing to the data for polar regions not being considered. It was recognized, and stated, that polar data might modify the results of that paper, but the nature and paramount importance of the corrections to be introduced were not foreseen.

17. The complete system of atmospheric currents is the combination of those

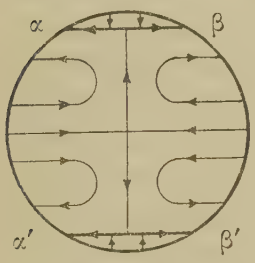


Fig. 7.

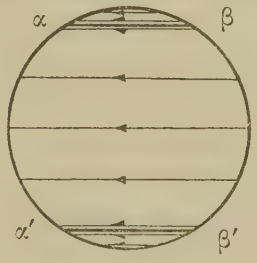


Fig. 8.

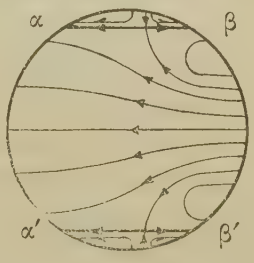
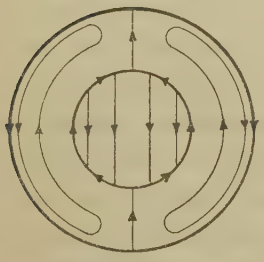
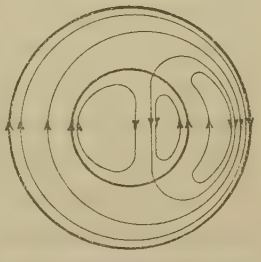


Fig. 9.





For the sake of clearness the angular distance of the auroral zone from the pole is shown as nearly 30° instead of about 23°.

The upper figures represent the elevation of the current systems as seen from the Sun. The lower figures are the plans of the upper ones in figs. 7 and 9.

illustrated in figs. 7, 8, but its precise form will naturally depend on the relative intensities of these two partial current systems, which are not indicated in the diagrams. Where the currents in the separate diagrams are in the same direction, the resultant current will naturally be in that direction also ; but where the separate currents are in opposite directions, it is essential to know their relative magnitudes in order to determine the direction of the resultant current : cases of the latter kind occur along the zones $\alpha\beta$, $\alpha'\beta'$, in the right (or post-meridiem) hemisphere as viewed from the sun, and along the equator

in the left (or ante-meridiem) hemisphere. It seems probable that at the equator the intensity of the storm-time currents exceeds that of the diurnal system, because the average storm-time change in the horizontal force exceeds the semi-range of the disturbance-diurnal variation of horizontal force at the equator (*e.g.*, for all-minus-quiet days the semi-range of the latter is about 4γ—*cf.* fig. 1, curve 5*c*—while the reduction of the daily mean horizontal force is about 10γ—*cf.* § 4). Hence the direction of the combined current flow will be the same all round the equator (as in fig. 8), but the intensity will be least at 6 h. and greatest at 18 h. Along the zones $\alpha\beta$, $\alpha'\beta'$, on the other hand, the current intensity in the diurnal system (fig. 7) seems to exceed the average, judging by the same criterion; for the semi-range of the diurnal variation (all-minus-quiet days) at Bossekop, for example, is over 50γ, while the all-minus quiet-day difference between the daily means is less than 15γ. Hence the resultant current along the zones is unlikely to be in the same direction all round. The general form of the resultant current system is shown in fig. 9.

The above criterion for judging the relative intensities of the two current systems is, perhaps, not completely satisfactory, because it assumes that the current intensity is proportional to the corresponding part of the horizontal force in the immediate neighbourhood; this is not exactly true, since the horizontal force at any point depends in varying degree upon the current-distribution at a distance as well as near at hand. Moreover, it depends on currents within the earth as well as upon the atmospheric currents, which alone are under consideration at present. If the relative intensities of the two parts of the horizontal force variations were nearly equal, these circumstances might render the application of the criterion unsafe; but in actual fact they differ so greatly, both near the equator and near the zones, that the conclusions thus drawn seem likely to be valid.

The Origin of the Electric Currents.

18. In §§ 10–17 and figs. 7–9 the attempt has been made to infer the general distribution of the external current systems responsible for the additional magnetic field observed during disturbance—it being recognized, however, that part of the field is due to secondary currents induced in the earth. Figs. 7–9 are confessedly somewhat idealized, and, in any case, are provisional and subject to modification (not, however, likely to be of a radical nature, at least as regards figs. 7, 8) when the present lack of polar data is supplied. The question now arises as to the origin of these atmospheric electric currents.

In the first place, it seems not unlikely that part of the external currents

are themselves secondary induced currents, the actual primary currents being confined to high latitudes (though doubtless extending further towards the equator at times of intense than at times of small disturbance). There is good ground for the belief that there are regions of the atmosphere whose conductivities are of independent origin and vary in different ways (the regions are not necessarily quite separate, but may merge into one another near their boundaries); the arguments in favour of this conclusion are based on the different modes of variation, with respect to sunspot epoch and magnetic activity, of the solar and lunar diurnal variations.* Over the region from the equator to beyond $\pm$ 60° latitude there is a conducting layer whose conductivity does not vary in unison with the intensity of magnetic disturbance, though it shows a large and regular variation throughout the sunspot cycle. In polar latitudes, however, there is a region whose conductivity increases with the intensity of magnetic disturbance and of auroral activity. It seems reasonable to suppose that these latter variations of conductivity, themselves due to varying intensity of precipitation of solar corpuscles into the atmosphere, largely govern the intensity of magnetic disturbance, and that the disturbance mainly originates in this region, where the magnetic variations are known to be particularly intense and variable with locality. No hypothesis as to how these primary currents are produced will be made at this stage, but the magnetic data give clear evidence of the existence of intense currents in the polar atmosphere. Such currents must necessarily induce secondary currents in any other conducting regions, whether in the earth or atmosphere. Mr. T. T. Whitehead and I have shown that in lower latitudes the field of the induced currents† may be more important than the direct field (in these latitudes) of the primary currents.

The study of these induced currents unfortunately involves very laborious numerical computations, and is not sufficiently advanced to permit of a precise separation of the current system in fig. 9 into primary and secondary parts. Until this can be attempted it is difficult to progress further towards a detailed theory of the origin of the primary electric current system. The theory outlined in my former paper on magnetic storms (*loc. cit.*, from p. 76, § 9 onwards) is inadequate to explain the wider range of facts here considered, even apart from the valid objections made by Prof. Lindemann against some of the numerical developments of that theory.‡

* *Cf.* a lecture on "Some Problems of Terrestrial Magnetism," 'J. Lond. Math. Soc.,' April, p. 131.f. (1927).

† "The Influence of Electromagnetic Induction within the Earth upon Terrestrial Magnetic Storms" : 'Trans. Internat. Math. Congress, Toronto,' vol. ii (1924).

‡ F. A. Lindemann, 'Phil. Mag.,' vol. 38, p. 669 (1919).

It was there supposed that the magnetic disturbance field is produced by electro-magnetic induction in the atmosphere, by means of mass motion of air in the upper regions, across the earth's permanent magnetic field ; that is, in the same way, essentially, as in the case of the solar and lunar diurnal magnetic variations, according to the generally accepted theory. This view differs from that advocated (particularly) by Birkeland, which attributes magnetic disturbance to the direct magnetic field of streams of charges moving freely outside, or at high levels in, the atmosphere. Schuster* has shown, however, that Birkeland's hypothesis is untenable owing, among other reasons, to the inability of streams, charged with sufficient intensity to produce the observed magnetic changes, to maintain themselves against the tendency to dispersion by the mutual electrostatic repulsion of the charges.

While retaining this basic feature of my former theory, the hypothesis there suggested as to the type of atmospheric motion involved must, I think, be abandoned. The motion was supposed to be vertical, first downwards and then, during the greater part of the storm, upwards. These movements were attributed to the initial impact of corpuscles entering the atmosphere from outside, and to the subsequent upward expansion caused by the mutual repulsion of the accumulated corpuscles, which were supposed electrically charged, and predominantly of one sign. Such motions would account for electromotive forces of the type required to produce the current system of which part is illustrated in fig. 7 of that paper. The conductivity of the layer in which the currents flow was also attributed to the ionizing effect of the corpuscles.

The more complete current system illustrated in figs. 7–9 of this paper (or the primary part of this system) cannot be explained so well by such vertical motions. Vertical motions of the supposed kind must, indeed, occur, but their velocity and energy do not seem adequate to give rise to the observed magnetic fields. It is more likely that the atmospheric motions, whatever their type, are always present, and with them the corresponding electromotive forces : but that the intensity of the currents which they produce is dependent on the degree of ionization and conductivity of the air.

The type and velocity of the atmospheric motions must be inferred from the magnetic data, so long as no independent evidence of such upper atmospheric movements is available. It seems premature to attempt any detailed inferences, but the motions in question are more likely to be horizontal than vertical, on general dynamical grounds ; moreover, near the poles, where the magnetic force is nearly vertical, horizontal movements of the air are more

* A. Schuster, ' Roy. Soc. Proc.,' A, vol. lxxxv, p. 44 (especially § 6) (1911).

effective than vertical movements in inducing electromotive forces. A general poleward drift of the air in auroral latitudes would produce westerly electromotive forces of the type required in fig. 7 ; but the composite current system of fig. 9 seems to require the motion in auroral latitudes to be diurnally periodic, *i.e.*, polewards at 6 h. and towards the equator at 18 h. A motion having a velocity potential involving the spherical harmonic factor P_1^1 with maxima and minima of pressure at 6 h. and 18 h. respectively would accord with fig. 9, provided the atmospheric conductivity in the auroral region is greater at 6 h. than at 18 h. Such a motion is consistent with the observations of the barometer at ground-level, though the 24-hourly component of the daily barometric variation is much affected by local circumstances ; but it is doubtful whether this oscillation of type P_1^1 extends without change of phase and (proportional) amplitude up to great heights in the atmosphere. It is by no means impossible that an independent diurnal circulation of the upper atmosphere exists, due to the absorption of ultra-violet radiation, or of corpuscular radiation, or to both these causes independently ; but at least until the separate primary and secondary parts of the disturbance field have been inferred with the aid of the mathematical investigations now in progress, it would be premature to discuss further the causes of the atmospheric current systems in the polar regions.

In conclusion, I wish to express my thanks to Mr. A. E. Ludlam for drawing the diagrams for this paper, and to the Government Grant Committee of the Royal Society for assistance towards the expense of some of the computations involved in the paper.

The Method of Images in Some Problems of Surface Waves.

By T. H. Havelock, F.R.S.

(Received May 26, 1926.)

Introduction.

1. When a circular cylinder is submerged in a uniform stream, the surface elevation may be calculated, to a first approximation, by a method due originally to Lamb for this case, and later extended to bodies of more general form : the method consists in replacing the cylinder by the equivalent doublet at its centre and then finding the fluid motion due to this doublet. In discussing the problem some years ago,* I remarked that if the solution so obtained were interpreted in terms of an image system of sources, we should then be able to proceed to further approximations by the method of successive images, taking images alternately in the surface of the submerged body and in the free surface of the stream. This is effected in the following paper for two-dimensional fluid motion, and the method is applied to the circular cylinder. It provides, theoretically at least, a process for obtaining any required degree of approximation, but, of course, the expressions soon become very complicated. It is, however, of interest to examine some cases numerically so as to obtain some idea of the degree of approximation of the first stage.

An expression is first obtained for the velocity potential of the fluid motion due to a doublet at a given depth below the surface of a stream, the doublet being of given magnitude with its axis in any direction. A transformation of this expression then gives a simple interpretation in terms of an image system. This system consists of a certain isolated doublet at the image point above the free surface, together with a line distribution of doublets on a horizontal line to the rear of this point ; the moment per unit length of the line distribution is constant, but the direction of the axis rotates as we pass along the line, the period of a revolution being equal to the wave-length of surface waves for the velocity of the stream. The contribution of each part of the image system to the surface disturbance is indicated.

Before proceeding to the circular cylinder, two cases are worked out in some detail, namely, a horizontal doublet and a vertical doublet. To a first approximation these give the surface disturbance of a stream of finite depth with an obstruction in the bed of the stream ; in the first case the bed of the stream is plane with a semi-circular ridge, and in the second case it has a more com-

* 'Roy. Soc. Proc.,' A, vol. 93, p. 524 (1917).

plicated form. Numerical calculations are made for both these cases, and graphs of the surface elevation are shown in figs. 1 and 2.

The second approximation for the circular cylinder is then investigated. The first stage is the surface effect due to a doublet at the centre, and the second is that due to a distribution of doublets on a certain semicircle. Expressions can be obtained for the complete surface elevation, but the calculations are limited to that part which consists of regular waves to the rear of the cylinder. The integrals are investigated and reduced to a form which permits of numerical evaluation. Calculations are carried out for various velocities for two different cases, namely, when the depth of the centre is twice, and three times, the radius. The results are tabulated for comparison, and one may estimate from these rather extreme cases the degree of approximation of the first stage. The effect of the second stage is to alter both the amplitude and the phase of the regular waves. The amplitude of the first-stage waves has a maximum for the velocity $\sqrt{(gf)}$, where f is the depth of the centre. It appears that the second stage increases the amplitude of the waves for velocities less than $\sqrt{(gf)}$ and decreases it for velocities above this value; further, the crests of the waves are moved slightly to the rear by an amount which varies with the speed. Some other possible applications of the method of images may be mentioned. For a doublet in a stream of finite depth, we can take successive images in the bed of the stream and in the free surface, and so build up the image system of a doubly infinite series of isolated doublets and of line distributions of doublets; this solution may be compared with the direct solution in finite terms which may be obtained in this case. Further, similar methods may be used for the three-dimensional fluid motion due to a doublet in a stream, and application made to the corresponding problem of a submerged sphere.

Image of Doublet in Stream.

2. We may either consider the doublet to be at rest in a uniform stream or to be moving with uniform velocity in a fluid otherwise at rest; we choose the latter alternative. Take Ox horizontal and in the undisturbed surface of the liquid, and Oy vertically upwards. Let the axes be moving with uniform velocity c in the direction of Ox, and let there be a two-dimensional doublet of moment M at the point $(0, -f)$ with its axis making an angle α with the positive direction of Ox. The velocity potential of the doublet is given by the real part of

$$\frac{\mathrm{M}e^{i\alpha}}{x+i(y+f)}. \qquad (1)$$

In order to keep the various integrals convergent and so to obtain a definite result, we adopt the usual device of a small frictional force proportional to velocity and in the limit make the frictional coefficient μ' tend to zero ; further, we neglect the square of the fluid velocity at the free surface.

If η is the surface elevation, the pressure equation gives the condition at the free surface,

$$\frac{\partial \phi}{\partial t} - g\eta + \mu'\phi = \text{const.}, \tag{2}$$

we have also, at the free surface,

$$\frac{\partial \eta}{\partial t} = -\frac{\partial \phi}{\partial y}. \tag{3}$$

And as we are dealing with the fluid motion which has attained a steady state relative to the moving axes, these conditions give, in terms of the velocity potential,

$$\frac{\partial^2 \phi}{\partial x^2} + \kappa_0 \frac{\partial \phi}{\partial y} - \mu \frac{\partial \phi}{\partial x} = 0, \tag{4}$$

to be satisfied at $y = 0$. Here we have put $\kappa_0 = g/c^2$ and $\eta = \mu'/c$.

We now assume the solution to be given by

$$\phi = -i\mathrm{M}e^{i\alpha} \int_0^\infty e^{i\kappa x - \kappa(y+f)}\, d\kappa + \int_0^\infty \mathrm{F}(\kappa)\, e^{i\kappa x + \kappa y}\, d\kappa. \tag{5}$$

The first term represents the doublet (1) in an equivalent form, valid for $y + f > 0$. The function $\mathrm{F}(\kappa)$ can now be determined by means of (4), and this gives

$$\mathrm{F}(\kappa) = i\mathrm{M}e^{i\alpha}\left(1 + \frac{2\kappa_0}{\kappa - \kappa_0 + i\mu}\right)e^{-\kappa f}. \tag{6}$$

Hence the velocity potential of the image system is

$$i\mathrm{M}e^{i\alpha} \int_0^\infty e^{i\kappa x - \kappa(f-y)}\, d\kappa + 2i\kappa_0 \mathrm{M}e^{i\alpha} \int_0^\infty \frac{e^{i\kappa x - \kappa(f-y)}}{\kappa - \kappa_0 + i\mu}\, d\kappa. \tag{7}$$

By comparison with (1) and the first term in (5), it is easily seen that the first term in (7) is the velocity potential in the liquid due to an isolated doublet at the image point $(0, f)$, of moment M with its axis making an angle $\pi - \alpha$ with Ox.

To interpret the second term in (7) we put

$$\frac{i}{\kappa - \kappa_0 + i\mu} = \int_0^\infty e^{-\mu p + i(\kappa - \kappa_0)p}\, dp, \qquad \mu > 0. \tag{8}$$

We then interchange the order of integration with regard to κ and p, and integrate first with respect to κ. The second term of (7) thus becomes

$$2i\kappa_0 \mathrm{M} e^{i\alpha} \int_0^\infty \frac{e^{-\mu p - i\kappa_0 p}}{x + p + i(f - y)}\, dp, \tag{9}$$

with $f - y > 0$.

By a comparison with (1), we see that the real part of (9) is the velocity potential of a line distribution of doublets along the line $y = f$, extending over the negative half of that line. The magnitude of the moment per unit length at the point $(-p, f)$ is $2\kappa_0 \mathrm{M} e^{-\mu p}$, and the axis at that point makes with Ox an angle $\kappa_0 p - \alpha - \frac{1}{2}\pi$.

It is necessary to retain the quantity μ while manipulating the integrals, but we may put it zero ultimately and we have the following result:—The image system of the doublet M at an angle α to Ox and at depth f below the surface consists of a doublet M at the image point at height f above the surface with the axis making an angle $\pi - \alpha$ with Ox, together with a line distribution of doublets to the rear of the image point of constant line density $2\kappa_0 \mathrm{M}$ and with the axis at a distance p in the rear making a positive angle $\kappa_0 p - \alpha$ with the downward-drawn vertical.

It is of interest to note how the parts of the image system contribute to the surface elevation. From the preceding equations we obtain

$$c\eta = \frac{2\mathrm{M}(f\cos\alpha - x\sin\alpha)}{x^2 + f^2} + 2\kappa_0 \mathrm{M} e^{i\alpha} \int_0^\infty \frac{e^{i\kappa x - \kappa f}}{\kappa - \kappa_0 + i\mu}\, d\kappa, \tag{10}$$

where the real part of the second term is to be taken.

The integral in (10) is transformed by contour integration, treating x positive and x negative separately; when μ is made zero ultimately, the complete expressions are

$$c\eta = \frac{2\mathrm{M}(f\cos\alpha - x\sin\alpha)}{x^2 + f^2} + 2\kappa_0 \mathrm{M} \int_0^\infty \frac{m\cos(mf - \alpha) - \kappa_0 \sin(mf - \alpha)}{m^2 + \kappa_0^2}\, e^{-mx}\, dm,$$

for $x > 0$; and

$$c\eta = \frac{2\mathrm{M}(f\cos\alpha - x\sin\alpha)}{x^2 + f^2} + 4\pi\kappa_0 \mathrm{M} e^{-\kappa_0 f} \sin(\kappa_0 x + \alpha) + 2\kappa_0 \mathrm{M} \int_0^\infty \frac{m\cos(mf + \alpha) - \kappa_0 \sin(mf + \alpha)}{m^2 + \kappa_0^2}\, e^{mx}\, dm, \tag{11}$$

for $x < 0$.

The first term in each case represents that part of the local surface disturbance due to the doublet and the isolated image doublet. The remaining terms

are due to the semi-infinite train of doublets behind the image point. Part of the effect is the train of regular waves to the rear of the origin, evidently associated with the periodicity in the direction of the doublets along the line distribution; and there is also a further contribution to the local surface disturbance, which we may regard as arising from the fact that the line distribution is semi-infinite and has a definite front.

Horizontal and Vertical Doublets.

3. With the axis of the doublet horizontal, we have the well-known first approximation to the submerged circular cylinder of radius a, if we take $\mathrm{M} = ca^2$. From (11), the surface elevation can be expressed in the form

$$\eta = \frac{2a^2 f}{x^2 + f^2} + 2a^2\kappa_0 \mathrm{P}, \quad x > 0,$$

$$\eta = \frac{2a^2 f}{x^2 + f^2} + 2a^2\kappa_0 \mathrm{P} + 4\pi\kappa_0 a^2 e^{-\kappa_0 f} \sin \kappa_0 x, \quad x < 0, \tag{12}$$

where P is the real part, for $x > 0$, of the integral

$$\int_0^\infty \frac{e^{-(x+if)m}}{m + i\kappa_0}\, dm. \tag{13}$$

Taking the axis of the doublet to be vertically upwards, we have $\alpha = \pi/2$ in the general formulæ; and, putting $\mathrm{M} = ca^2$ in this case also, we obtain

$$\eta = -\frac{2a^2 x}{x^2 + f^2} - 2a^2\kappa_0 \mathrm{Q}, \quad x > 0,$$

$$\eta = -\frac{2a^2 x}{x^2 + f^2} + 2a^2\kappa_0 \mathrm{Q} + 4\pi\kappa_0 a^2 e^{-\kappa_0 f} \cos \kappa_0 x, \quad x < 0, \tag{14}$$

where Q is the imaginary part of the integral (13). This integral may be expressed formally in terms of $li\,(e^{f-ix})$, where li denotes the logarithmic integral, and may be expanded in various forms. For the numerical calculations which follow, it was found simplest to use the series

$$\int_0^\infty \frac{e^{-(\alpha+i\beta)u}}{u + i}\, du = -(\mathrm{A} + i\mathrm{B})\, e^{-(\beta - i\alpha)},$$

$$\mathrm{A} = \gamma + \log r + \sum_1^\infty \frac{r^n}{n!} \cos n\theta,$$

$$\mathrm{B} = \pi - \theta - \sum_1^\infty \frac{r^n}{n!} \sin n\theta, \tag{15}$$

where

$$r = (\alpha^2 + \beta^2)^{\frac{1}{2}}, \qquad \tan\theta = \alpha/\beta, \qquad \text{and} \qquad \gamma = 0{\cdot}57721.$$

The series is sufficiently simple for calculation, though in some of the cases it was necessary to take a large number of terms.

For both the horizontal and vertical doublets we take

$$\mathrm{M} = ca^2, \qquad f = 2a, \qquad \kappa_0 f = 4. \tag{16}$$

This means that we take the velocity to be such that the wave-length of the regular waves is $\frac{1}{2}\pi f$. We are assuming, in each case, a given doublet at depth f below the surface of deep water. The only restrictions so far are the general ones due to neglecting the square of the fluid velocity at the free surface, and the consequent limitation to waves of small height. From this point of view the data of (16) are rather extreme; but, this being understood, it may be permissible to use them for a comparison of the two cases. With the values in (16), the calculations are comparatively simple, and lead to graphs which can be drawn suitably on the same scale throughout; these are shown in figs. 1 and 2, where the unit of length is the quantity a.

In fig. 1, there is a horizontal doublet at C; the arrow shows the direction

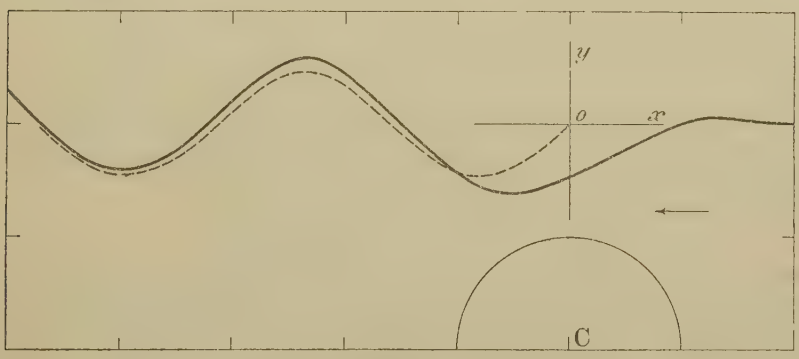


Fig. 1.

of the stream assuming the doublet to be stationary, and Ox is in the undisturbed surface. The surface elevation was calculated from (12) for the case (16). The broken curve shows the regular sine waves to which the disturbance approximates as we pass to the rear. This solution is also the first approximation for a submerged cylinder of radius a; or, again, to the same order, it gives the effect caused by a semicircular ridge on the bed of a stream of depth twice the radius. From this point of view the diagram may be compared with that given by Kelvin* for a small obstruction on the bed of a stream of finite depth.

* Kelvin, 'Math. and Phys. Papers,' vol. 4, p. 295.

Fig. 2 shows the corresponding curves for a vertical doublet, calculated from (14) for the case (16); the doublet is at the point C. Here, again, the broken curve shows the cosine term of the solution to which the disturbance approximates.

We may also regard this as an approximate solution for the flow of a stream

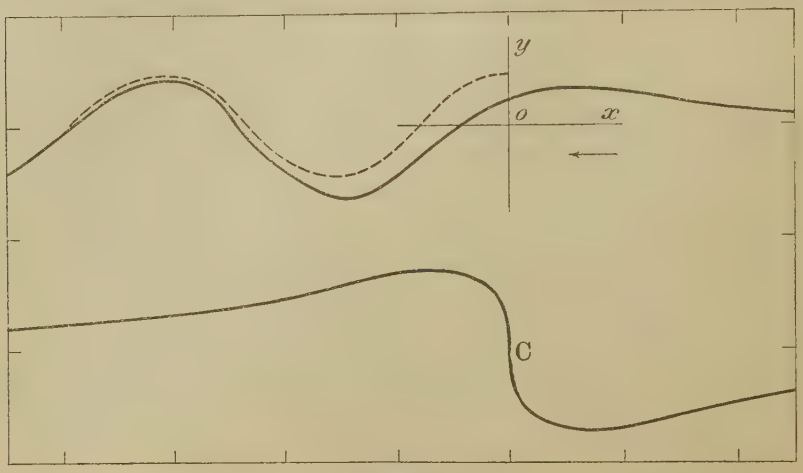


Fig. 2.

over a bed of a certain form. This is obtained by taking the zero stream-line for the combination of the uniform stream and a vertical doublet at C under the conditions given in (16); the equation of this curve is

$$(y + 2a)\{x^2 + (y + 2a)^2\} + a^2x = 0, \qquad (17)$$

and its form is shown in the figure. Fig. 2 may be compared with a graph given by Wien* for the case of a sudden small rise in the bed of a stream.

It is interesting to note the general similarity of the surface elevation in the two cases shown in figs. 1 and 2; although the regular waves are given by a sine curve in one case and a cosine curve in the other, that is only because of the different position of the origin relative to the general form of the obstacle.

Second Approximation for Circular Cylinder.

4. We may now carry out further approximations for a circular cylinder in a uniform stream by the method of successive images. Reference may be made to fig. 3, which is not drawn exactly to scale.

The image of the stream in the circle is a horizontal doublet M at the centre C. The image of M in the free surface is a doublet — M at the image point C_1

* W. Wien, 'Hydrodynamik,' p. 206.

together with a trail of doublets to the rear of C_1. The image of this system in the circle gives a doublet $-Ma^2/4f^2$ at C_2, together with a certain line distribution of doublets on the semicircle on CC_2. So the process could be carried on, but we shall stop at this stage.

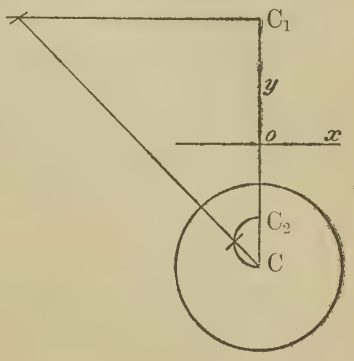


Fig. 3.

From the results already given, we could build up complete expressions for the velocity potential and surface elevation for each stage. It would be of interest to work these out graphically to compare with fig. 1; but the expressions soon become complicated and their evaluation difficult, especially for the immediate vicinity of the origin. We shall therefore limit the study to the regular waves established in the rear of the cylinder. We have seen that the regular waves of the first approximation, due to the doublet ca^2 at C, are given by

$$\eta = 4\pi\kappa_0 a^2 e^{-\kappa_0 f} \sin \kappa_0 x\,; \qquad x<0. \tag{18}$$

We take the next stage in two parts. First we have an isolated horizontal doublet of moment $-ca^4/4f^2$ at C_2, whose co-ordinates are $(0, -f+a^2/2f)$. From (11) it follows that the contribution of this doublet to the regular waves is

$$\eta = -\pi\kappa_0 a^4 f^{-2} e^{-\kappa_0 (f - a^2/2f)} \sin \kappa_0 x\,; \qquad x<0. \tag{19}$$

Next we consider the line distribution of doublets to the rear of C_1 and its image in the circle. Referring to the results in § 2, there is at the point $(-p, f)$ an elementary doublet of moment $2\kappa_0 ca^2\,dp$, with its axis making an angle $\kappa_0 p - \frac{1}{2}\pi$ with the positive direction of Ox. The image of this in the circle is a doublet at the point whose co-ordinates are

$$-\frac{a^2 p}{p^2+4f^2}, \quad -f+\frac{2a^2 f}{p^2+4f^2}; \tag{20}$$

the moment of the doublet is $2\kappa_0 ca^4 \,.\, dp/(p^2+4f^2)$, and its axis makes with Ox the angle

$$2\tan^{-1}(p/2f) - \kappa_0 p + \tfrac{1}{2}\pi. \tag{21}$$

From (11) we can now write down the waves due to this doublet. It should be noted that the expression will hold for

$$x + \frac{a^2 p}{p^2+4f^2} < 0.$$

If, therefore, we wish to obtain the complete expression for this part of the surface elevation at a point in the range $-a^2/4f < x < 0$, we should have to integrate with respect to p between appropriate variable limits. We shall consider only points to the rear of this range, so that the limits for p are 0 and ∞. This being understood, the distribution of doublets on the semicircle CC_2 contributes to the regular waves a part given by

$$\eta = 8\pi\kappa_0{}^2 a^4 e^{-\kappa_0 f} \int_0^\infty e^{\frac{2a^2\kappa_0 f}{p^2+4f^2}} \cos\left\{\kappa_0\left(x + \frac{a^2 p}{p^2+4f^2}\right) + 2\tan^{-1}\frac{p}{2f} - \kappa_0 p\right\} \frac{dp}{p^2+4f^2}. \quad (22)$$

Putting $p = 2f\tan\frac{1}{2}\theta$, this becomes

$$\eta = 2\pi\kappa_0{}^2 a^4 f^{-1} e^{-\kappa_0 f + \kappa_0 a^2/4f}(\mathrm{A}\cos\kappa_0 x - \mathrm{B}\sin\kappa_0 x), \quad (23)$$

where

$$\mathrm{A} = \int_0^\pi e^{h\cos\theta}\cos(\theta + h\sin\theta - k\tan\tfrac{1}{2}\theta)\,d\theta,$$

$$\mathrm{B} = \int_0^\pi e^{h\cos\theta}\sin(\theta + h\sin\theta - k\tan\tfrac{1}{2}\theta)\,d\theta,$$

with $h = \kappa_0 a^2/4f$ and $k = 2\kappa_0 f$.

5. In the applications to be made, h and k are positive, h is less than unity and is usually a small fraction. In these circumstances, the integrals may be evaluated by expansion in power series of h. It can be shown, after a little reduction, that we have

$$\mathrm{A} = 2\sum_0^\infty \frac{h^n}{n!}\mathrm{L}_{n+1}; \qquad \mathrm{B} = 2\sum_0^\infty \frac{h^n}{n!}\mathrm{M}_{n+1}; \quad (24)$$

where

$$\mathrm{L}_r = \int_0^{\pi/2}\cos(2r\phi - k\tan\phi)\,d\phi$$

$$\mathrm{M}_r = \int_0^{\pi/2}\sin(2r\phi - k\tan\phi)\,d\phi. \quad (25)$$

The quantities L and M may be evaluated in terms of known functions by a reduction formula. It can readily be shown that

$$(r+1)\mathrm{L}_{r+1} = k\mathrm{L}_r'' - k\mathrm{L}_r' + r\mathrm{L}_r, \quad (26)$$

the accents denoting differentiation with respect to k; or denoting this operation by D, we have

$$r!\,\mathrm{L}_r = (k\mathrm{D}^2 - k\mathrm{D} - r - 1)(k\mathrm{D}^2 - k\mathrm{D} - r - 2)\ldots(k\mathrm{D}^2 - k\mathrm{D})\mathrm{L}_0. \quad (27)$$

The quantity M satisfies similar relations.

Further, we have

$$\mathrm{L}_0 = \int_0^{\pi/2} \cos(k \tan\phi)\, d\phi = \tfrac{1}{2}\pi e^{-k}$$

$$\mathrm{M}_0 = -\int_0^{\pi/2} \sin(k \tan\phi)\, d\phi = -\tfrac{1}{2}\{e^{-k} li(e^k) - e^k li(e^{-k})\}. \qquad (28)$$

We shall find it necessary to go as far as the sixth term in numerical calculation of A and B; we therefore record to this order explicit expressions for L and M obtained from (27) and (28).

$$\mathrm{L}_1 = \pi k e^{-k},$$
$$\mathrm{L}_2 = -\pi k(1-k)e^{-k},$$
$$\mathrm{L}_3 = \tfrac{1}{3}\pi k(3-6k+2k^2)e^{-k},$$
$$\mathrm{L}_4 = -\tfrac{1}{3}\pi k(3-9k+6k^2-k^3)e^{-k},$$
$$\mathrm{L}_5 = \tfrac{1}{15}\pi k(15-60k+60k^2-20k^3+2k^4)e^{-k},$$
$$\mathrm{L}_6 = -\tfrac{1}{45}\pi k(45-225k+300k^2-150k^3+30k^4-2k^5)e^{-k},$$
$$\mathrm{M}_1 = -ke^{-k} li(e^k) + 1,$$
$$\mathrm{M}_2 = k(1-k)e^{-k} li(e^k) + k,$$
$$\mathrm{M}_3 = -\tfrac{1}{3}k(3-6k+2k^2)e^{-k} li(e^k) + \tfrac{1}{3}(1-4k+2k^2),$$
$$\mathrm{M}_4 = \tfrac{1}{3}k(3-9k+6k^2-k^3)e^{-k} li(e^k) + \tfrac{1}{3}k(5-5k+k^2),$$
$$\mathrm{M}_5 = -\tfrac{1}{15}k(15-60k+60k^2-20k^3+2k^4)e^{-k} li(e^k) + \tfrac{1}{15}(3-28k+44k^2-18k^3+2k^4),$$
$$\mathrm{M}_6 = \tfrac{1}{45}k(45-225k+300k^2-150k^3+30k^4-2k^5)e^{-k} li(e^k) + \tfrac{1}{45}k(93-198k+124k^2-28k^3+2k^4).$$

6. The first case we shall examine is that already discussed in § 3, a cylinder whose centre is at a depth of twice the radius. It has been remarked that this is an extreme case, but it has the advantage, as far as the calculations are concerned, of magnifying the difference between the first and second approximations and so of lightening the numerical work involved. In the notation of the previous sections, we have

$$f = 2a; \quad k = 2\kappa_0 f = 4\pi f/\lambda_0; \quad h = \kappa_0 a^2/4f = k/32. \qquad (29)$$

Collecting the terms in (18), (19) and (23), the regular waves established to the rear of the cylinder are given by

$$\eta/a = \pi k e^{-\frac{1}{2}k}\sin\kappa_0 x - \tfrac{1}{16}\pi k e^{-\frac{7}{16}k}\sin\kappa_0 x + \tfrac{1}{16}\pi k^2 e^{-\frac{15}{32}k}(\mathrm{A}\cos\kappa_0 x - \mathrm{B}\sin\kappa_0 x). \qquad (30)$$

The first term is the first approximation, and the amplitude in this case has a maximum at $k = 2$, or when the velocity is such that the wave-length is $2\pi f$.

We shall calculate the value of (30) for k equal to 10, 8, 6, 4, 2, 1 and 0·5, given in order of increasing velocity. Omitting the intermediary steps for the numerical values of the L and M functions, the following table gives the values of A and B, calculated from (24), for these values of k and with $h = k/32$ in each case.

k	10	8	6	4	2	1	0·5
A	0·021	0·064	0·204	0·646	1·805	2·311	1·891
B	−0·418	−0·522	−0·716	−0·950	−0·596	0·668	1·742

The simplest form in which to show the difference made by the second approximation is to express (30) in each case in the form

$$\eta/a = \mathrm{D} \sin \kappa_0 (x + \xi), \tag{31}$$

and compare it with the first approximation

$$\eta/a = \mathrm{C} \sin \kappa_0 x. \tag{32}$$

A comparison of D and C gives the alteration in the amplitude of the waves ; further, there is an alteration in phase expressed as a displacement of the crests to the rear by an amount ξ.

In this form the final numerical values, for $f = 2a$, are given in the following table :—

$c/\sqrt{(ga)}$.	C	D	ξ/a.
0·63	0·212	0·263	0·006
0·71	0·460	0·568	0·017
0·82	0·939	1·159	0·050
1·0	1·701	2·046	0·148
1·41	2·312	2·396	0·468
2·0	1·906	1·721	0·669
2·83	1·223	1·081	0·595

We see that the second approximation makes a considerable difference in the amplitude in this case ; but it should be noted that, in addition to the depth being only twice the radius, the velocities are relatively large, the wave-length at the lowest velocity being about $1\frac{1}{4}$ times the depth.

The amplitude C has a maximum at the speed $\sqrt{(2ga)}$; and it appears from the table that the second approximation increases the amplitude below this velocity and diminishes it at higher velocities. It seems that the rearward displacement, given by ξ, also has a maximum, amounting to about two-thirds of the radius of the cylinder.

7. It is clear, from the form of the expressions for the surface elevation, that the accuracy of the first approximation increases rapidly as the depth of the cylinder is increased or as we take relatively smaller velocities. Without pursuing the calculations in this direction, we shall take one other case which is not quite so extreme as in the previous section. We take the depth of the centre to be three times the radius ; the data are now

$$f = 3a\,;\ k = 2\kappa_0 f\,;\ h = \kappa_0 a^2/4f = k/72. \tag{33}$$

In this case, instead of (30), we have

$$\eta/a = \tfrac{2}{3}\pi k e^{-\frac{1}{12}k} \sin \kappa_0 x - \tfrac{1}{5\,4}\pi k e^{-\frac{17}{12}k} \sin \kappa_0 x + \tfrac{1}{5\,4}\pi k^2 e^{-\frac{25}{12}k} (\mathrm{A} \cos \kappa_0 x - \mathrm{B} \sin \kappa_0 x). \tag{34}$$

The following table shows the values of A and B, with $h = k/72$, calculated for convenience at the same values of k as before :—

k	10	8	6	4	2	1	0·5
A	0·008	0·033	0·137	0·540	1·747	2·311	1·898
B	−0·324	−0·428	−0·626	−0·911	−0·644	0·663	1·732

With the same notation as in (31) and (32), the results are given in the following table :—

$c/\sqrt{(ga)}$.	C	D	ξ/a.
0·77	0·141	0·149	0·001
0·87	0·307	0·329	0·006
1·0	0·626	0·677	0·024
1·22	1·134	1·222	0·088
1·73	1·541	1·558	0·295
2·45	1·270	1·214	0·409
3·46	0·816	0·802	0·324

The calculations were made for the same values of k; and as we have taken $\sqrt{(ga)}$ as the unit of velocity, we get a different set of velocities, but they cover much the same range. We notice that the decrease of the ratio a/f from $\frac{1}{2}$ to $\frac{1}{3}$ has diminished considerably the difference between C and D, and also the displacement ξ. The results have the same general character as we noted in the previous case.

In any given case there are two significant quantities involved: one is the ratio of the radius to the depth and the other is the ratio of the wave-length to the depth. It would require a more elaborate numerical study than has been attempted here to enable us to state precisely the degree of accuracy of the first approximation for given values of these ratios.

The Emission of Soft X-Rays by Different Elements.

By O. W. Richardson, F.R.S., Yarrow Research Professor of the Royal Society, and F. S. Robertson, M.I.E.E., Senior Lecturer in Electrical Engineering, King's College, London.

(Received May 9, 1927.)

In a paper by Richardson and Chalkin* it was shown that the efficiency of different elements as emitters of soft X-rays increased with increasing atomic number of the respective elements far less rapidly than does the corresponding property for ordinary X-rays. Whereas the efficiency for the latter is proportional to the atomic number, it was found that the efficiency for soft X-rays was roughly proportional to the square roots of the atomic numbers for the elements tested. The experiments were confined to the four elements carbon (At. No. 6), iron (26), nickel (28) and tungsten (74), and were spread over a long period of time, during which various changes were made in the apparatus. They were neither primarily intended nor well adapted to investigate with any accuracy the comparative efficiency of different sources of the rays. Accordingly, the apparatus described in this paper has been constructed. By means of it, six elements can be tested in quick succession under similar conditions without opening up the apparatus. We have tested the 14 elements tabulated below. These are all that we have been able to obtain which are sufficiently

* 'Roy. Soc. Proc.,' A, vol. 110, p. 273 (1926).

refractory to stand purification by high vacuum bombardment and which we have been able, by one means or another, to form into plates or sheets of the requisite size.

As a result of these experiments it is quite clear that *the efficiency for soft X-ray emission is a periodic function of the atomic number*, like other phenomena which depend on the superficial electronic distribution.

The apparatus used in these experiments consisted of a quartz containing vessel, the details of which are shown in fig. 1. In this were mounted two filaments, the structure supporting the targets, a pair of nickel condenser plates, a nickel photoelectric plate around which was an electrode of the same metal to catch the electrons emitted from it, and an earth shield, also of the same metal.

It was so arranged that six targets could be experimented upon each under precisely the same conditions without opening the tube. The dimensions of each target were 2 cm. by 1 cm. by 0·1 cm.; they were held in a lantern-like structure made of molybdenum, details of which are shown in fig. 2. The targets were all interchangeable and were fitted in position between parallel molybdenum rods 0·1 cm. diameter, by means of vertical grooves cut in their edges to fit the rods; a hexagonal molybdenum cap fitted down against shoulders on the top ends of the rods and held the whole in position.

The lower hexagonal plate of the lantern-like structure was drilled in the centre and riveted against a shoulder on the top of a central nickel rod which supported it. The lower end of this rod was drilled out to a fairly good fit on a long brass rod fixed to a stopper at the bottom of the quartz apparatus; this formed an axis about which the whole structure could be turned so as to present each target in turn to the working filament. Six equally spaced grooves in the lower part of the nickel rod and a pin in the side of the brass rod made it possible for one to be quite sure that each target in succession could be placed in precisely the same position relative to the working filament and the other parts of the apparatus. The targets could be easily moved around by means of an external electromagnet, which acted upon two short soft iron rods fixed to the movable structure as shown.

By removing the stopper at the bottom of the apparatus the whole structure supporting the targets could be taken out, fitted with new targets as desired, and put back again without disturbing either of the filaments.

The vessel was very carefully exhausted and its contents freed as far as possible from gas by an efficient mercury vapour pump suitably backed. During this process the vessel itself was heated externally and the targets were

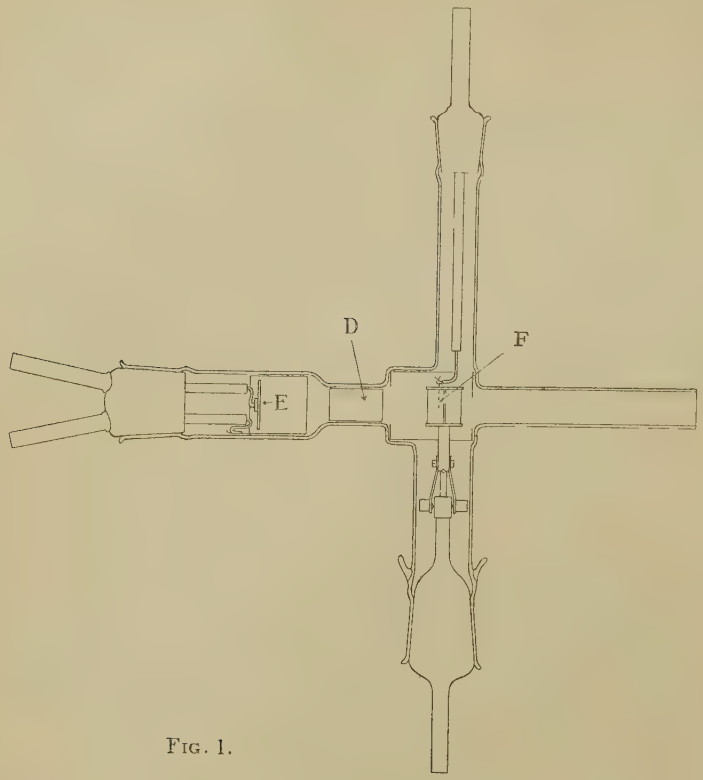
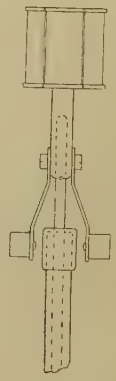


Fig. 1.



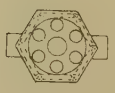


Fig. 2.

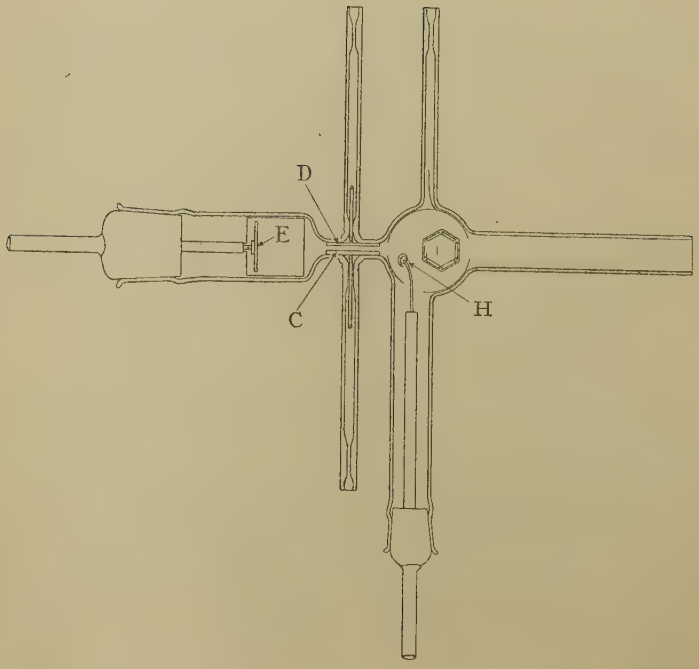


Fig. 1.—Apparatus.

Fig. 2.—Target and Holder.

repeatedly raised to a red heat by bombarding them at 550 volts from the filament F mounted inside the ring of targets as shown in fig. 1. The filament F was used for this purpose only, and being placed inside the ring of targets it did not contaminate the working faces by the tungsten vapour given off at the high temperature necessary for bombardment.

All the observations recorded in this paper were taken with the same working filament at H, which, after having been once thoroughly cleaned up, was always used at a comparatively low temperature and did not seem to produce any appreciable contamination of the working faces of the targets.

Two liquid air traps were employed and the McLeod gauge, used for rough measurement of the pressure, together with the final glass tap, were inserted between these. By this arrangement, the mercury vapour from the McLeod gauge, together with any vapours from tap grease, etc., were effectively cut off from the apparatus by a liquid air trap containing charcoal. The pressure in the apparatus itself was estimated from time to time by altering the electrical connections in such a way as to convert it into a sensitive ionisation gauge.

A diagram of the electrical connections used is shown in fig. 3. The value of

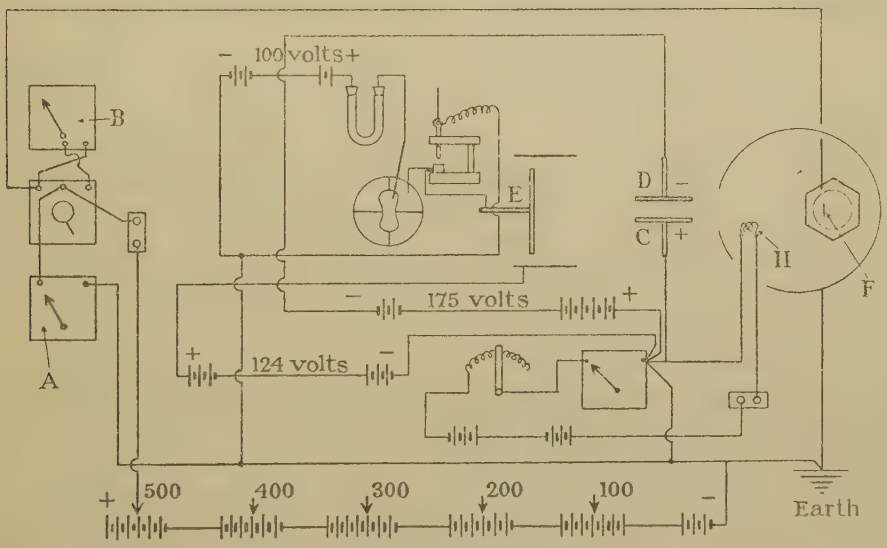


Fig. 3.—Electrical Connections.

the target voltage used was measured on the voltmeter A and was the voltage between the negative end of the filament and the target. The thermionic current passing to the target was measured by the microamperemeter B.

The X-radiation was recorded by means of the photoelectric current produced when the beam of X-radiation fell upon a nickel plate connected to the insulated

quadrants of a Dolezalek electrometer, the rate of change of potential and the capacity of the system being measured in the usual way. The beam of X-radiation was made to pass between two condenser plates C and D before impinging on the nickel plate, the condenser plate C was at earth potential and D at 175 volts negative to earth. By a preliminary experiment it was made clear that no measurable current due to electrons or positively or negatively charged ions could pass through the condenser thus charged to the photo-electric plate E.

Targets were prepared from samples of 14 different elements; and as only six could be placed in the apparatus at any one time, they were grouped as follows:—

1st group.		2nd group.		3rd group.	
Carbon	6	Chromium	24	Silicon	14
Iron	26	Manganese	25	Copper	29
Nickel	28	Iron	26	Molybdenum	42
Molybdenum	42	Cobalt	27	Palladium	46
Tungsten	74	Nickel	28	Silver	47
Platinum	78	Copper	29	Gold	79

The numbers following the names of the elements are their respective atomic numbers.

It will be noticed that group 2 and group 3 each contained two targets which had previously been experimented upon; this was done in order that they could be used in comparing one set of experiments with another taken perhaps some weeks afterwards, it having been necessary in the meantime to open up the apparatus and re-exhaust it.

In each experiment the thermionic current producing the X-radiation was measured and also the photoelectric current produced by a definite portion of that X-radiation when it fell upon a nickel plate.

The ratio $\frac{\text{Photoelectric Current}}{\text{Thermionic Current}}\left(\frac{ip}{it}\right)$ was noted and is shown plotted against Atomic Number for typical sets of observations picked from many which were taken on each group.

Figs. 4, 5 and 6 show the results obtained with 300, 400 and 500 volts potential difference exciting the X-rays respectively. The general nature of the results for the different elements is very much the same for each voltage, the main difference being that the ratio ip/it increases, for each element, at the higher voltages and, in fact, for a given element it is approximately proportional

to the voltage. In a given group of the periodic table the ratio is highest near the middle of the group and falls fairly steadily as the beginning of the next group is approached. This happens for each of the groups Cr to Cu, Mo to Ag and W to Au, and the values for the elements C and Si, both of which are at the middle of a group of the periodic table, seem high. Thus, the efficiency found for Si is higher than that of any of the elements tested, except Cr and Mo,

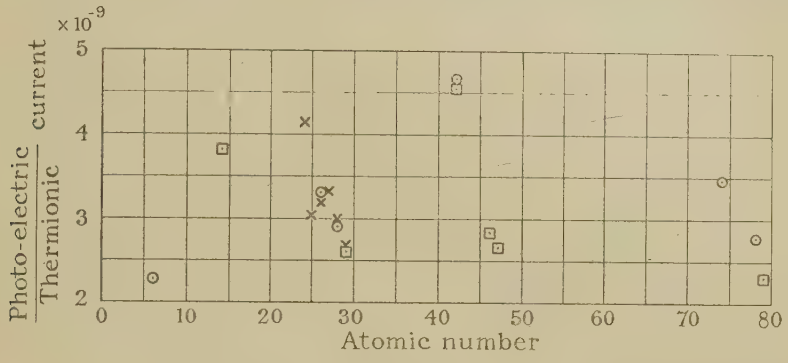


FIG. 4.

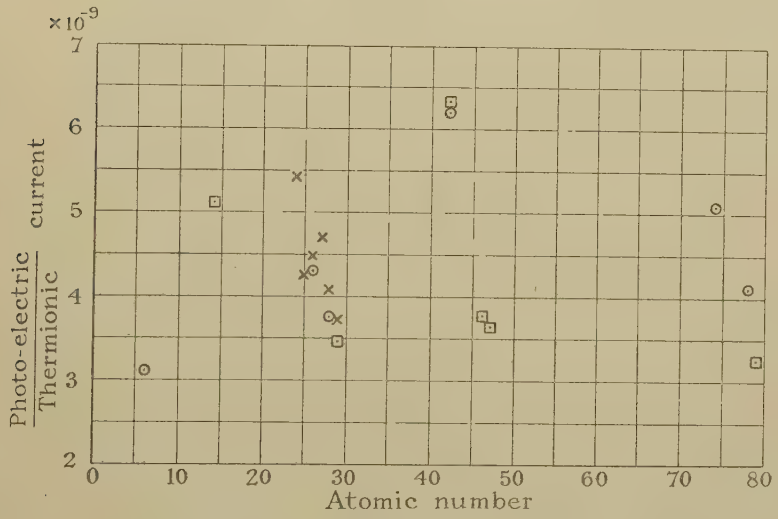


FIG. 5.

and although the value found for C is low absolutely, it is nevertheless practically the same as that found for Au, the element of highest atomic weight of any which have been tested.

The only group for which we could obtain a run of consecutive elements is that of Cr, Mn, Fe, Co, Ni and Cu. The results for these at 100, 200, 300, 400 and 500 volts are shown in fig. 7. It appears that the fall in efficiency as the

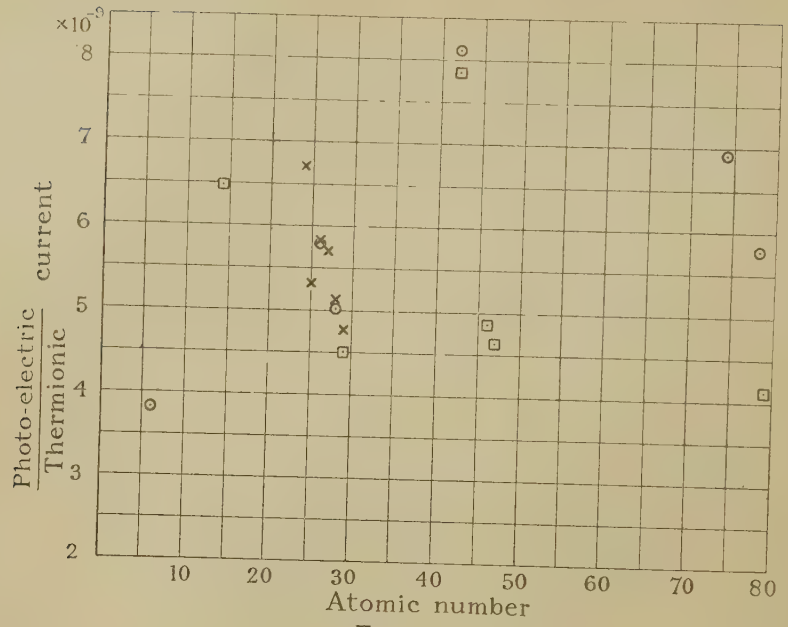

×10⁻⁹
8
7
6
5
4
3
2
Photo-electric current / Thermionic current
10
20
30
40
50
60
70
80
Atomic number

FIG. 6.

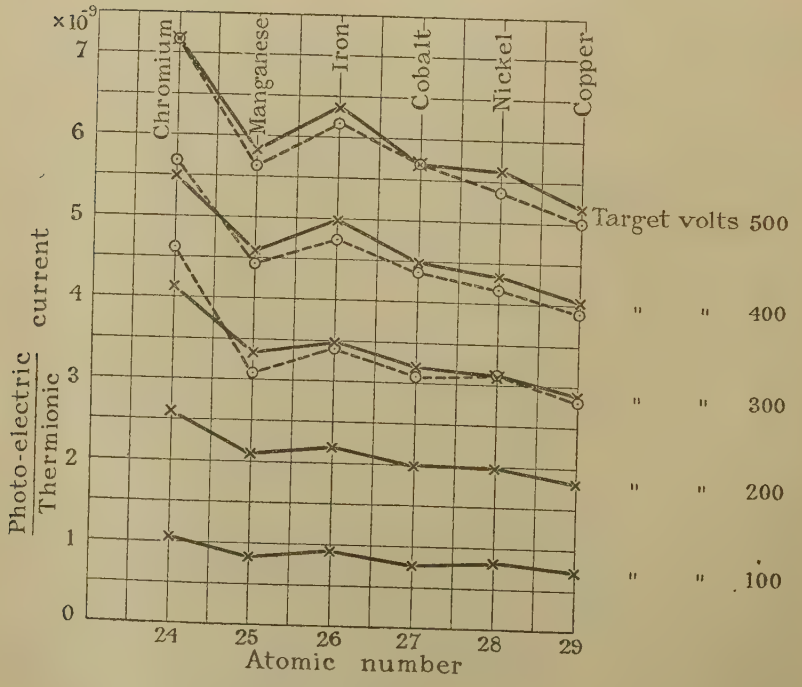

×10⁻⁹
7
6
5
4
3
2
1
0
Photo-electric current / Thermionic current
Chromium
Manganese
Iron
Cobalt
Nickel
Copper
Target volts 500
" " 400
" " 300
" " 200
" " 100
24
25
26
27
28
29
Atomic number

FIG. 7.

atomic number increases is not a steady one, but one which oscillates. Thus the ratio for manganese was always found to be less than that for iron, and the value for cobalt was less than the average of the values for iron and nickel.

The present results partly crystallise and partly modify the conclusions drawn by Richardson and Chalkin. The inference that the efficiency is proportional to the square root of the atomic number of the elements was evidently based on inadequate data and must be abandoned. The conception of the phenomena as dominated, to a first approximation, by a reduced "effective atomic number" may still be regarded as valid, with the effective atomic number a periodic function of the actual atomic number. If we denote the effective atomic number by Ne, and, as on page 278 of the paper by Richardson and Chalkin, carry out a computation of the product αNe, where α is the number of electrons ejected from the nickel photoelectric detecting plate per unit quantum of soft X-radiation incident on it, we obtain the numbers in the last row of the following table :—

Table 1.

Element	C	Si	Cr	Mn	Fe	Co	Ni	Cu	Mo	Pd	Ag	W	Pt	Au
Atomic No.	6	14	24	25	26	27	28	29	42	46	47	74	78	79
$\frac{\alpha Ne}{2\beta}$	3·54	5·92	6·42	4·70	5·32	5·03	4·85	4·10	7·16	4·58	4·15	5·36	5·08	3·57

The values of αNe (here designated by $\alpha Ne/2\beta$, see below) for C, Fe, Ni and W are appreciably higher than those found by Richardson and Chalkin, viz., for C, 0·87, for Fe (2·32), for Ni 2·11 and W 3·47. There seems no doubt that the present values are the right ones. Similar values were obtained on a number of occasions after various changes in the apparatus. The pressure was measured by the ionisation gauge method and was always less than 3×10^{-7} mm. The difference may be due in part to the fact that Richardson and Chalkin used a copper photoelectric target, whereas we used a nickel one, but preliminary tests (unpublished) by Miss L. P. Davies have shown that there is no radical difference between the results from copper and nickel photoelectric targets in this field. We were prevented from making tests with a copper photoelectric target in our apparatus owing to an accident. Another likely reason for part of this discrepancy is that in the present experiments the rays were investigated in the direction of the normal to the target instead of at 45° to it.

If we write the photoelectric current ip as

$$ip = \alpha ne, \qquad (1)$$

where " n " is the number of quanta of radiation incident on the target in unit time and " e " the electronic charge, and assume that the average value of the energy of a quantum of the X-radiation is βeV, where " V " is the primary voltage on the tube and β is approximately independent of V, the energy of the X-radiation which gives rise to the photoelectric current is

$$n\beta e\mathrm{V} = \beta/\alpha \,.\, ip\mathrm{V}. \qquad (2)$$

If the electronic current to the anticathode is ie, the energy liberated per electron impact on the target is

$$\frac{4\pi \mathrm{A}d\,(d-x)}{h^2b^2}\,\frac{\beta}{\alpha}\,\frac{ip}{ie}\,e\mathrm{V}. \qquad (3)$$

Here " A " is the area of the effective part of the anticathode, " h " the distance of the condenser plates apart, " b " their breadth, " d " the distance of the anticathode from the farthest, and " x " from the nearest, end of the condenser plates.* A number of values of ip/ie divided by " V " (in electrostatic units) are given in Table II.

Table II.—Values of $ip/ie \div$ V in E.S. Units at Different Voltages.

Date.	At 100 volts.	At 200 volts.	At 300 volts.	At 400 volts.	At 500 volts.
		Carbon (6)			
B. February 5th	—	—	2·288	2·348	2·31
A. ,, 5th	2·685	2·457	2·451	2·45	2·432
A. ,, 11th	2·49	2·408	2·387	2·284	2·513
		Silicon (14).			
March 28th....	—	—	3·827	3·865	3·893
		Chromium (24).			
A. March 11th....	3·157	3·895	—	4·27	4·302
B. ,, 11th	—	—	4·124	4·125	4·30
,, 16th	—	—	4·15	4·061	4·027
		Manganese (25).			
March 8th	—	—	3·064	—	—
A. ,, 11th	2·517	3·138	3·098	3·339	3·384
B. ,, 11th	—	—	3·324	3·42	3·571
,, 16th	—	—	3·038	3·188	3·198
		Iron (26).			
A. February 5th	3·10	2·89	3·265	3·13	3·44
B. ,, 5th	—	—	3·32	3·24	3·48
A. ,, 11th	2·865	2·865	3·155	—	—
A. March 11th....	2.718	—	—	—	—
,, 16th....	—	—	3·186	3·36	3·493

* For details, see Richardson and Chalkin, *loc. cit.*, p. 274.

Table II.—continued.

	Date.	At 100 volts.	At 200 volts.	At 300 volts.	At 400 volts.	At 500 volts.
			Cobalt (27).			
	March 8th....	—	—	3·00	—	—
A.	,, 11th....	2·300	2·992	3·10	3·286	3·41
B.	,, 11th....	—	—	3·20	3·36	3·42
	,, 16th....	—	—	3·334	3·524	3·426
			Nickel (28).			
A.	February 5th	2·454	2·375	2·60	2·963	3·212
B.	,, 5th	—	—	2·917	2·828	3·019
A.	,, 11th	2·418	2·521	2·786	—	—
B.	,, 11th	—	—	2·645	2·972	3·14
	March 8th....	—	—	2·830	—	—
A.	,, 11th	2·472	—	—	—	—
	,, 16th	—	—	2·993	3·072	3·078
			Copper (29).			
	March 8th....	—	—	2·603	—	—
A.	,, 11th	2·217	2·6 (approx.)	—	—	—
	,, 16th	—	—	2·683	2·807	2·87
	,, 28th	—	—	2·619	2·622	2·708
			Molybdenum (42).			
A.	February 5th	3·783	4·066	4·512	—	—
B.	,, 5th	—	—	4·69	4·665	4·865
A.	,, 11th	4·008	4·148	4·555	—	—
	March 28th....	—	—	4·58	4·768	4·721
			Palladium (46).			
	March 28th....	—	—	2·856	2·854	2·926
			Silver (47).			
	March 28th....	—	—	2·686	2·76	2·79
			Tungsten (74).			
A.	February 5th	3·366	3·322	3·522	3·945	4·241
B.	,, 5th	—	—	3·463	3·823	4·150
A.	,, 11th	3·666	3·307	3·496	3·88	4·221
			Platinum (78).			
A.	February 5th	2·853	2·832	2·918	3·200	3·351
B.	,, 5th	—	—	2·789	3·096	3·472
A.	,, 11th	3·144	3·056	3·12	3·343	3·517
			Gold (79).			
	March 28th....	—	—	2·31	2·463	2·477

The figures opposite the same date *and* the same letter A or B, where such a letter is assigned, correspond to the same set of readings in which there was relatively little change in the working conditions of the apparatus. They are more reliable for comparing the different elements than figures selected indiscriminately. On the other hand, the numbers as a whole are thought to give a fair idea of the reliability of the absolute values. The single sets of numbers assigned to silicon, palladium, silver and gold do not rest alone, but are supported by other measures which give substantially the same values in terms of molybdenum and copper, but values which are about 12 per cent. higher for all six elements. These absolute numbers are affected by an error and have not been included.

The numbers for $ip/ie \div V$ are not very different for the different voltages, showing that β is approximately independent of "V." The further elucidation of the data is impossible with any certainty in the present state of the subject owing to our ignorance of the quantities α and β. Some experiments made about two years ago by one of us (O.W.R.), in collaboration with Mr. A. A. Newbold, indicated that most of the photoelectrons had an amount of energy much less than that which corresponded to the primary voltage. These results were not published as it was not found possible to eliminate satisfactorily the complications arising from secondary and reflected radiations which are present with the stopping potential method which was used. This ambiguity has been overcome by Mr. E. Rudberg* by using a magnetic deflection method of analysing the velocity distribution among the ejected photoelectrons. In this way he has been able to avoid the troublesome secondary effects, and the results show that with a carbon anticathode and a copper photoelectric detector some 70 per cent. of the photoelectrons have energies not exceeding about 10 volts with 700 volts on the tube. This result may mean either that (1) the X-ray quantum generated at the anticathode may have an amount of energy comparable with that corresponding to the voltage on the tube, but that, owing to various possible transformations, this may give rise to the emission, not of one electron with the corresponding energy, but to a number of electrons with lower energies at the photoelectric surface. Simons† has shown that something of this kind occurs with ordinary X-rays. In this case the value of α might be greater than unity. Or (2), the X-rays generated at the anticathode may have a low average frequency compared with eV/h, where "V" is the tube voltage. This would correspond to a value of 2β less than unity. Of course, both complications may occur with the possibility $\alpha > 1 > 2\beta$.

The only direct way to unravel these factors would appear to involve grating methods, and the prospect of success with them at present is not encouraging. In any event it is interesting to observe from Table 1 that if $\alpha/2\beta$ is of the order unity, the values of Ne (the reduced value of the nuclear charge which is necessary if Kramer's theory of X-ray generation is to apply to these soft X-rays) are similar to the numbers of electrons in the various atoms which have negative energies comparable with the energies of the electrons falling on the anticathode.

We wish to acknowledge our indebtedness to Prof. W. T. Gordon, who very kindly cut some of the more difficult specimens for us.

* 'Nature,' p. 704, May 14th, 1927.

† L. Simons, 'Proc. Phys. Soc. London,' vol. 37, p. 58 (1925).

On the Nature of Wireless Signal Variations—I.

By E. V. Appleton, M.A., D.Sc., F.R.S., Wheatstone Professor of Physics, King's College, London, and J. A. Ratcliffe, B.A., Clerk Maxwell Student, Cavendish Laboratory, Cambridge.

(Received May 30, 1927.)

[Plate 5.]

1. *Introduction.*

In a recent communication* two experimental methods of investigating the atmospheric deflection of electric waves have been described. In one of these methods interference phenomena, which take the form of a succession of signal maxima and minima at the receiving station, are artificially produced by a small continuous change of transmitter wave-length. From the study of these phenomena information relating to the path difference and relative intensities of the ground and atmospheric waves may be deduced. In the second type of experiment the angle of incidence of the down-coming waves is deduced from an examination of the electric and magnetic vectors in the stationary interference system produced at the ground by the waves incident from above and reflected at the earth's surface. In a more recent communication† some further results, obtained by means of the " wave-length change " method, have been described, which indicate that the variations of signal intensity produced at moderate distances from a continuous-wave transmitting station are due to the variable nature of the rays returned from the upper atmosphere, and not to slight variations of transmitter wave-length. In this second communication, however, certain points were left unelucidated. For example, it was pointed out that the wave-length change experiments gave us no information as to the relative importance of the parts played by changes in intensity, phase polarization, or angle of incidence of the downcoming waves in producing the ordinary nocturnal signal variations.

The development of the second type of experiment mentioned above (in which the electric and magnetic vectors of the stationary interference system produced at the ground are examined) has yielded considerable information on these and allied points, and it is with these developments that the present communication deals. The refinements in the technique of the original

* ' Roy. Soc. Proc.,' A, vol. 109, p. 621 (1925).

† ' Roy. Soc. Proc.,' A, vol. 113, p. 450 (1926).

experiments may be grouped under three heads. In the first place, we have employed automatic photographic registration for recording the galvanometer deflections corresponding to the magnitudes of various vectors. Such registration is specially useful when simultaneous records of the variations of any two vectors are required.

Secondly, by using a combination of a loop and vertical antenna, such as is used in commercial practice for a different purpose,* it has been found possible to eliminate the effects of the ground ray at the receiving station. With such a system it is possible to receive only the rays deviated by the upper atmosphere and to investigate them independently. For convenience, such an arrangement has been named the "suppressed ground-ray system."

Thirdly, the study of natural "fading" has been supplemented by experiments in which fading is artificially produced by small continuous changes of transmitter wave-length. In such experiments the signal variations are definitely due to changes in the phase-difference between ground and atmospheric waves, and not to changes of relative intensity. In this way a difficulty, which is inherent in the study of natural fading by means of the ordinary signal-strength measurements, is removed, for with natural signal variations we are, in general, unable to distinguish between changes brought about by variations of phase and by variations of intensity.

The experiments to be described fall roughly into two series. The first series is concerned with the determination of the angle of incidence of the downcoming waves, and the second with the direct investigation of these waves by means of the suppressed ground-ray system. These are described in Parts I and II respectively. To prevent repetition, the general experimental details relating to both series of experiments and the theoretical ideas underlying them have been grouped together in Sections 2, 3 and 4 of Part I. The bearing of the results obtained on the general problem of the nature of fading is discussed mainly in Section 2 (*c*) of Part II.

2. *Factors Governing Choice of Experimental Conditions.*

The earlier experiments, in which the electric and magnetic forces due to the downcoming waves were compared, were made at Cambridge, using the signals from the London (2LO) transmitter. In that case it was found that the effect of the atmospheric waves relative to that of the ground wave was small. Observations on signals received from more distant transmitters

* Such a combination is used for wireless direction finding without bilateral ambiguity, since it possesses a polar reception sensitivity diagram of cardioid form.

showed that, as the distance of transmission was increased, the effect of the atmospheric waves, compared with that of the ground wave, increased, so that, at a certain distance, the vertical electric force produced by the downcoming waves was equal to that of the direct wave, resulting in a complete absence of signals when the two were out of phase. This critical distance of maximum signal variation is found to vary widely, indicating a somewhat large variation in the "reflection-coefficient" of the ionized layer, but for a wave-length of 400 metres it is roughly 100 to 150 miles.* In the present series of experiments it was considered desirable to work at such a distance that the ground ray was always the stronger, and yet such that marked changes of signal intensity were experienced. Since two transmitters (the London Broadcasting Station and the National Physical Laboratory transmitter) were available in London, the receiving station was moved from Cambridge to Peterborough, thus increasing the distance of transmission from 55 to 78 miles. This distance has proved satisfactory, in that only on rare occasions have the downcoming waves been sufficiently strong enough to interfere completely with the ground waves and cause the signal intensity to fall to zero.

3. *The Transmitting and Receiving Stations.*

Most of the observations have been made on the ordinary transmissions from the London (2LO) Broadcasting Station, but special early morning transmissions of unmodulated waves from this station have been used in experiments carried out during the sunrise period. The National Physical Laboratory transmitter (5HW) has been used for the specially controlled experiments, in which the wave-length is continuously varied. The wave lengths used have been 365 metres (2LO) and 400 metres (5HW). A reference to the shape of the aerial system of the 2LO transmitter will be made later.

The aerial systems of the receiving station at Peterborough were as follows :—

(*a*) A vertical aerial, 16 metres high, at the top of which was situated a symmetrical wire cage to increase the capacity to earth. This aerial system was used in conjunction with an earth consisting of buried metal plates ;

* Some other experiments indicate that this distance of maximum fading varies with wave-length, being greater the greater the wave-length. For example, maximum fading of the signals from Daventry (wave-length 1,600 metres) is found about 400 to 600 miles from the transmitter.

(*b*) a large loop aerial consisting of a single turn of wire* in the form of a triangle. This loop was situated in the vertical plane containing the transmitting and receiving stations ;

(*c*) a vertical loop aerial similar to (*b*), but situated with its plane at right angles to the direction of transmission.

The two loops were arranged with their vertical axes of symmetry coincident, and the vertical wire aerial was situated along this axis, as shown in fig. 1. In this arrangement there is no magnetic coupling between any two of the three aerial systems. To eliminate " antenna effect " when using the loops as tuned receivers, the circuits were completed as indicated in fig. 2. The loop

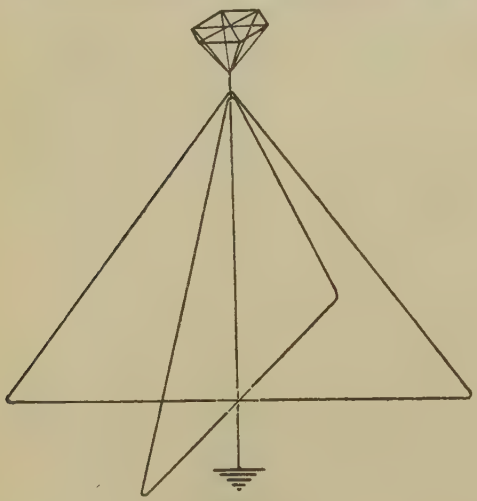


Fig. 1.

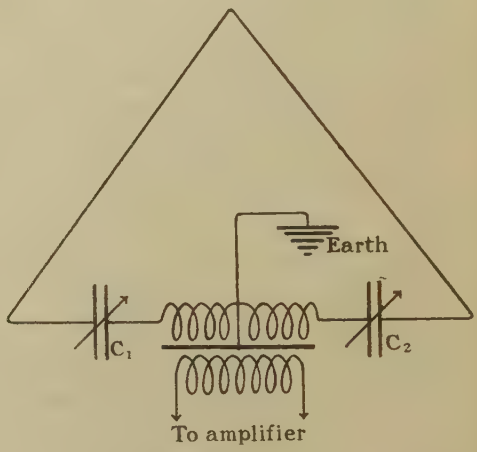


Fig. 2.

in each case was coupled to its amplifier by means of a transformer with the primary earthed at its mid-point. When the transformer was the only coupling between the aerial system and the amplifier it was shielded as shown. If the linkage was effected by way of a closed tuned circuit, the shielding was usually dispensed with. To keep conditions symmetrical with respect to the earthed point, the loop was tuned by means of two equal condensers, C_1 and C_2. When the loop was used in the untuned state the circuit was similar, except that the condensers were absent.

The following amplifiers and recording apparatus were used in conjunction with the aerial systems.

* A large single-turn loop was used in preference to a smaller loop of many turns because of the greater reception efficiency and comparative freedom from " antenna effect " of the former.

(*a*) For transmissions on a constant wave-length a two-stage neutrodyne high-frequency amplifier was used. To the anode circuit of the second stage of this amplifier was coupled a coil in series with a stable crystal detector and a reflecting moving-coil galvanometer. The galvanometer deflections were photographically recorded on a Cambridge Drum Camera. Records taken during the day-time showed that the amplifier was quite constant over the time (usually half-an-hour) required to obtain a photographic record of signal variations. Very frequently simultaneous records were made of the signals obtained on two-different aerial systems. In such cases it was found possible to prevent inter-action between the two amplifiers by enclosing one of them in a box lined with galvanised iron and earthing suitable points of the circuits. In each series of experiments a test was made to ensure that no inter-action was taking place. The deflections of the two galvanometers were recorded simultaneously on the same drum.

(*b*) For the special transmissions in which the transmitter wave-length was continuously varied through a small range a resistance was inserted in the aerial tuning circuit and a transformer-coupled high-frequency amplifier was used. This amplifier, the transformers of which were wound with resistance wire, was kindly loaned to us by Captain H. G. Round of the Marconi Company. The anode circuit of the last stage was coupled to a stable crystal detector and Einthoven string galvanometer. When it was necessary to know the law of overall-response of the aerial-amplifier-galvanometer assembly a calibration was effected by inserting electromotive forces of known relative magnitude into the aerial circuit.

4. *General Theoretical Considerations.*

The theoretical considerations underlying the study of downcoming waves by examining the stationary interference system at the ground have already been given,* and we need only recapitulate the results here. The relevant details are best illustrated by reference to fig. 3, which is drawn in the plane of propagation (*i.e.*, the vertical plane containing the transmitting and receiving stations).

Let O be the site of the receiving station at which a ground wave (electric and magnetic vectors E_0 and H_0 respectively) and a downcoming wave (incident at angle i_1) are received. The downcoming wave may be resolved

* Appleton and Barnett, 'Electrician,' vol. 95 p. 678 (1925).

into two components, one (E_1, H_1) with electric vector in the plane of propagation and the other (E_1', H_1') with electric vector perpendicular to this plane. We

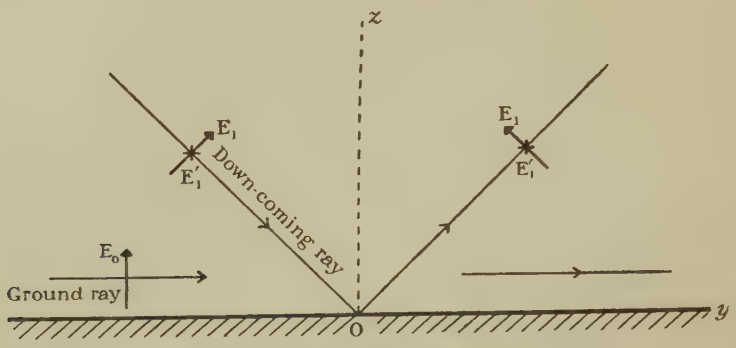


FIG. 3.—(The x axis is at right angles to the paper.)

shall call these, following Eckersley, the normally and abnormally polarized components respectively. Then the electric and magnetic forces at O may be written as follows :—

$$\left.\begin{aligned} E_x &= 0, \quad H_x = H_0 \sin pt + 2H_1 \sin (pt + \theta) \\ E_y &= 0, \quad H_y = 2H_1' \cos i_1 \sin (pt + \theta') \end{aligned}\right\}, \qquad (1)$$

$$E_z = E_0 \sin pt + 2E_1 \sin i_1 \sin (pt + \theta), \quad H_z = 0,$$

where p is the angular frequency of the waves and θ and θ' represent the phase differences between the ground wave and the normal and abnormal components of the downcoming wave respectively. The nature of the downcoming waves at any instant is, therefore, specified in terms of the quantities E_1 (or H_1), E_1' (or H_1'), i_1, θ and θ'. Now the three aerial systems which have been described above, namely vertical aerial, loop in the plane of propagation and loop at right angles to the plane of propagation, are acted on by electromotive forces which are proportional to the three vectors E_z, H_x and H_y respectively. It is, therefore, necessary to devise experimental methods which will yield most directly information relating to the variation of the quantities specifying the downcoming rays in terms of the observable signal magnitudes E_z, H_x and H_y.

In the experiments involving the simultaneous recording of the signals received on two aerials it is, of course, essential that the field acting on one of the aerials should not be distorted by the presence of the other. But the distortion of an incident field is merely the effect of the superposition on the incident field of the re-radiated field from the distorting antenna. We can, therefore, be certain that the field incident on one aerial is not distorted by the

presence of another aerial if we arrange, as in the present experiments, that the two aerial systems have no appreciable magnetic coupling.

5. *Measurement of the Angle of Incidence of the Downcoming Waves.*

Two methods have been employed for measuring the angle of incidence of the downcoming waves. The first (method (i)) is a direct development of that which has already been described.* The second (method (ii)) has been developed to supplement and also to overcome certain difficulties inherent in method (i). The two methods are described immediately below.

(*a*) *Method* (i).—In these experiments the vertical aerial and loop in the plane of propagation have been used. The amplifier coupled to the vertical aerial is acted on at the input end by an electromotive force of amplitude $K_A E_A$ where K_A is a constant depending on the aerial circuit and

$$E_A = E_z = \sqrt{E_0^2 + 4E_1^2 \sin^2 i_1 + 4E_0 E_1 \sin i_1 \cos \theta}. \qquad (2)$$

Now the amplitude of the electromotive force developed in the loop is proportional to $\frac{d}{dt}(H_x)$, but since in an electric wave the electric and magnetic forces are proportional,

$$\frac{d}{dt}(H_x) = H_0 p \cos pt + 2H_1 p \cos (pt + \theta) = k \,.\, p\,[E_0 \sin pt + 2E_1 \cos (pt + \theta)],$$

where k is a constant. The amplitude of the electromotive force impressed on the input end of the loop amplifier is therefore $K_L E_L$ where

$$E_L = \sqrt{E_0^2 + 4E_1^2 + 4E_0 E_1 \cos \theta}, \qquad (3)$$

and K_L is a constant the value of which is partly determined by the loop circuit constants.

Let us now consider the case in which the signal variations ΔE_A and ΔE_L due to the downcoming waves are small so that $E_0^2 \gg E_1^2$. As a first approximation, therefore, (2) and (3) become

$$\left.\begin{aligned} E_A &= E_0 + 2E_1 \sin i_1 \cos \theta \\ \text{and} \qquad E_L &= E_0 + 2E_1 \cos \theta \end{aligned}\right\}. \qquad (4)$$

Now the small variations ΔE_A and ΔE_L may be due to changes of intensity or

* 'Roy Soc. Proc.,' A, vol. 109, p. 621 (1925). See also Smith-Rose and Barfield, 'Roy. Soc. Proc.,' A, vol. 110, p. 580 (1926).

of phase of the downcoming wave, and if we suppose that, in the general case, both these changes are taking place, we have

$$\left.\begin{aligned}\Delta E_A &= \sin i_1 \,(2 \cos \theta \Delta E_1 + 2E_1 \Delta (\cos \theta)\,)\\ \Delta E_L &= 2 \cos \theta \Delta E_1 + 2E_1 \Delta (\cos \theta)\end{aligned}\right\}. \qquad (5)$$

Thus whether the small changes of received signal strength are due to changes of intensity or to changes of phase of the downcoming wave we may write

$$\frac{\Delta E_A}{\Delta E_L} = \sin i_1. \qquad (6)$$

In the present series of experiments the vertical aerial and the loop in the plane of propagation were arranged so that each could be switched on to either of two neutrodyned amplifiers, suitably shielded. The assemblies were adjusted so that, in the daytime, when the downcoming waves are of negligible intensity for the wave-lengths employed, the same galvanometer deflection was obtained when either amplifier was connected to the vertical aerial or to the loop. It may be shown* that, with this adjustment, the ratio of the departure ΔA of the night time signal current from the day-time value on the vertical aerial set, to the similar departure ΔL on the loop set is given by

$$\frac{\Delta A}{\Delta L} = \frac{\Delta E_A}{\Delta E_L} = \sin i_1. \qquad (7)$$

Simultaneous photographic records were made of the natural signal variations on the loop and on the vertical aerial. At intervals of about three minutes the aerial systems and amplifiers were interchanged.† Since it is essential that there should be no inter-action between the two systems, tests were made using the steady day-time signals to verify that the presence of a signal in one system produced no effect on the galvanometer deflection on the other.

(*b*) *Method* (ii).—As previously mentioned, method (i) is only applicable when the signal variations are small. It has only been possible to use it at Peterborough during the sunrise and sunset periods or in mid-night periods when the variations were abnormally small.

It will readily be seen from (2) and (3) that we may obtain similar results with large signal variations if it may be assumed that E_1 is constant and that the variations are due solely to changes in θ. Such conditions will obtain if

* *Vide* ' Roy. Soc. Proc.,' A, vol. 109, p. 632 (1925).

† A more detailed examination of the problem shows that any difference in the two amplifier characteristics may be allowed for by taking the geometric mean of the results obtained before and after interchanging the amplifiers.

we produce the signal variations by rapidly and continuously changing the transmitter wave-length through a small range. Such a wave-length change produces what may be termed "phase maxima and minima" in both aerial systems.

Now the electromotive forces acting on an amplifier connected to vertical aerial or loop are $K_A E_A$ and $K_L E_L$ respectively. At a phase maximum $\cos\theta$ is unity so that

$$\left.\begin{aligned} K_A E_A &= M_A = K_A (E_0 + 2E_1 \sin i_1) \\ \text{and} \\ K_L E_L &= M_L = K_L (E_0 + 2E_1) \end{aligned}\right\}, \tag{8}$$

while at a phase minimum

$$\left.\begin{aligned} K_A E_A &= m_A = K_A (E_0 - 2E_1 \sin i_1) \\ \text{and} \\ K_L E_L &= m_L = K_L (E_0 - 2E_1) \end{aligned}\right\}. \tag{9}$$

Thus

$$\frac{M_A - m_A}{M_A + m_A} \Big/ \frac{M_L - m_L}{M_L + m_L} = \sin i_1, \tag{10}$$

where

$$\frac{M_A - m_A}{M_A + m_A} \text{ and } \frac{M_L - m_L}{M_L + m_L}$$

are respectively the fractional changes in the input electromotive forces acting on the amplifiers connected to the vertical aerial and loop.

For determinations of the angle of incidence by this method the flatly-tuned amplifier, in conjunction with the Einthoven galvanometer, was arranged so that it could be switched alternately to the vertical aerial or to the loop circuits in rapid succession. Photographic records were made of special signals transmitted from the National Physical Laboratory. The usual type of transmission consisted of continuous changes of wave-length of 10 metres (on a mean wave-length of 400 metres) made in about five seconds, with intermediate steady dashes lasting about two seconds.* The amplifier was connected alternately to the loop and vertical aerial. When using this method the changes of electromotive force are not necessarily small, and they cannot be taken as proportional to the variations in galvanometer deflection. The galvanometer scale was, therefore, calibrated in terms of input electromotive force immediately after taking each record.

(*c*) *Experimental Results.*—The photographic records obtained by method (i) all showed in a striking way that the natural night-time signal variations are

* Telephonic communication between the transmitter and receiver was used for control purposes.

greater on a loop than on a vertical aerial. This is illustrated by a sample record shown in fig. 4 which shows simultaneous signal variations on both systems. The galvanometer deflections are measured in opposite directions. The zero line for each galvanometer is marked on the figure, that for the upper curve being at the top of the record and that for the lower curve being at the bottom. To test for any possible difference in the two amplifiers the sets were interchanged half-way through the record. Fig. 4 has been chosen to

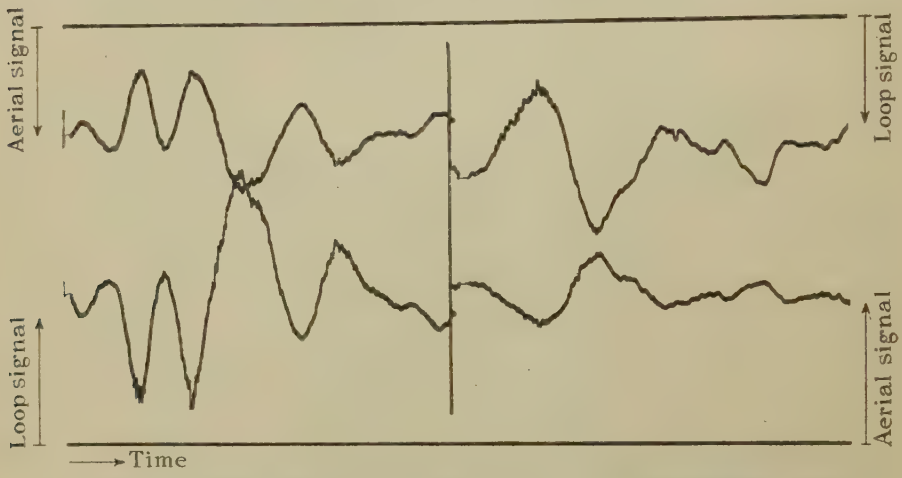


Fig. 4.—Signal Variations on Aerial and Loop, 7 p.m., September 29, 1926.

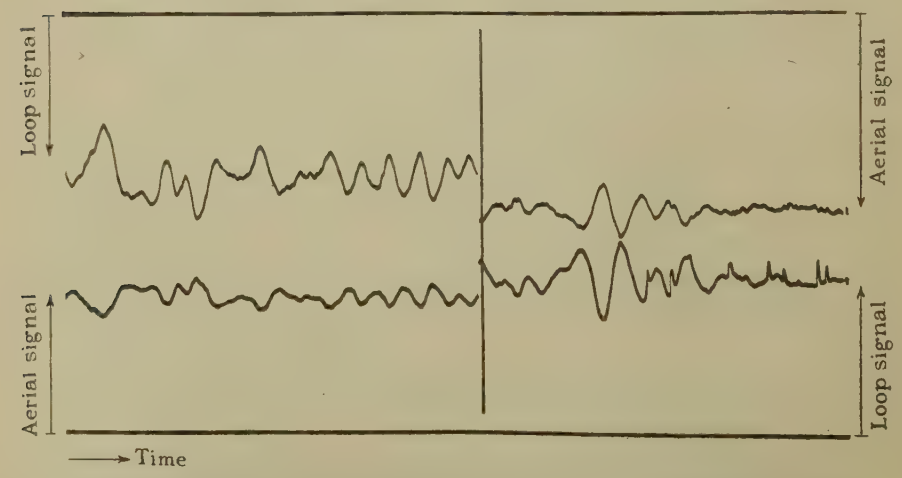


Fig. 5.—Signal Variations on Aerial and Loop, 6.30 a.m., September 30, 1926.

illustrate the relative values of loop and vertical aerial signal variations. It is an example of conditions when the variations are too large to admit of the

deduction of the angle of incidence of the downcoming waves by method (i). Fig. 5 shows the kind of record of similar type which is suitable for purposes of calculation. Each record lasted six minutes, the sets being interchanged after three minutes.

Plate 5A is a sample photographic record obtained by method (ii). It shows the signal variations due to a wave-length change at the transmitter, as the amplifier is switched alternately to the loop and vertical aerial. The record shows the signal current as a function of the time during six wave-length changes of 10 metres each, on a mean wave-length of 400 metres. There was a slight pause between each wave-length change which is indicated by the steady signal current. The speed of the camera paper was such as to make the record shown last 20 seconds. It is seen that the amplitude of the interference "fringes" is always greater for the loop than for the vertical aerial. It is interesting to note that records such as that of Plate 5A contain in themselves a proof of the atmospheric ionized layer theory which has been used to explain the nocturnal phenomena in wireless transmission. The fact that interference "fringes" are produced, due to a wave-length change at the transmitter, shows that there must be two transmission paths of different length, while the fact that the fringes are more pronounced on the loop aerial shows that the indirect waves must come down from above.

We now consider some numerical results which have been obtained from an analysis of many records similar to those of fig. 5 and Plate 5A. It is always found, by both methods, that the values of sin i_1, calculated from various parts of the same record, show considerable variations. The significance of this fact is considered below. Because of this somewhat wide variation in the angle of incidence we have, for comparison purposes, usually found the mean value of sin i_1 for the period occupied in making a photographic record.

Method (i) has been used in the evenings, one or two hours after sunset, and in the mornings, about sunrise.* The following are examples of mean values of sin i_1 for several $\frac{1}{4}$-hour records obtained on various evenings in September, 1926:

$$\sin i_1 = 0{\cdot}48\,;\ 0{\cdot}43\,;\ 0{\cdot}56\,;\ 0{\cdot}61.$$

These values show the kind of variation which is found in sin i_1 from night to night, and from time to time on the same night. The mean value of sin i_1 obtained from all the records made during September, 1926, is 0·56.

* Unfortunately, the applicability of both methods is limited. Method (i) may only be used when the signal variations are small (*e.g.* at sunrise and sunset), while method (ii), although suitable for large signal variations, may only be used when interference from Broadcasting stations (British and Continental) is negligible (*e.g.* in the early morning).

Method (ii) has generally been used in the early morning period, a few hours before sunrise. It has been found that the values of sin i_1 generally increase as sunrise is approached. Such a variation of the mean value of sin i_1, with time, is shown in Table I, which gives the results obtained on the morning of December 4, 1926.

Table I.

Time.	4.57 a.m.	5.08 a.m.	5.26 a.m.	5.33 a.m.	5.37 a.m.
Sin i_1	0·41	0·51	0·58	0·70	0·72

Variation of angle of incidence (measured by Method (ii)*) on the morning of December 4th,* 1926.

This increase of sin i_1, as sunrise approaches, suggests that the atmospheric deviating layer moves downwards, owing to the increased ionization produced by the sun's rays, and the evidence supports that which has been obtained from observations on the number of interference fringes produced by a wave-length change at the transmitter.* Since the smaller values of sin i_1 appear to be associated with night conditions, while the larger values are associated with the transitional conditions which obtain within one or two hours of sunrise, it appears reasonable to compare these larger "sunrise" values with the values obtained by method (i) in the transition period near sunset. The mean of such "sunrise" values of sin i_1, obtained by method (ii) on various mornings during November and December, 1926, is found to be 0·57.

In carrying out the determinations of sin i_1 by means of method (ii) we usually allowed one wave-length change on each photograph to be recorded entirely on either loop or vertical aerial, so that the total number of fringes for a given wave-length change could be counted. In this way it was possible, making certain assumptions, to calculate the path difference between the ground and atmospheric waves, and thus estimate the effective height at which the atmospheric deflection takes place. Now, if we make the same assumption of atmospheric deviation at a comparatively sharp boundary, we can also use the mean values of sin i_1, found from the same record, to estimate this height independently. The mean number of interference fringes produced by a continuous wave-length change from 395 ms. to 405 ms., during the "sunrise" period, is found to be 7·5. The effective height of the atmospheric

* 'Roy. Soc. Proc.,' A, vol. 113, p. 456 (1926).

deflecting layer corresponding to this is 104 kms. The effective height corresponding to sin $i_1 = 0{\cdot}56$ (the mean value by method (i)) is 92 kms., while that corresponding to sin $i_1 = 0{\cdot}57$ (the mean value by method (ii)) is 88 kms. While there seems to be definitely a difference between the heights as determined by the two methods (the possible significance of which is mentioned below), there can be little doubt as to the order of magnitude of the height of the deviating layer at the transitional periods near sunrise and sunset.

It has been mentioned that in the night periods, several hours before sunrise, abnormally small values of sin i_1 are sometimes observed. It is found that these small values are always accompanied by a correspondingly large number of "fringes" due to the wave-length change. This appears to indicate that they are due to a deviating layer at heights much greater than those mentioned above (about 250 kms.). In this case a comparison of the effective heights as obtained from the number of fringes and from the angle of incidence does not show a very close agreement. This is probably due to the fact that in assuming "reflection" to take place at the same height for all wave-lengths we neglect a certain correction, as has been previously pointed out.* Special interest is attached to this correction because its magnitude gives us information about the gradient of ionisation in the layer. By comparing the effective heights deduced from "fringe measurements" and from measurements of the angle of incidence, we hope to obtain information about this correction term.

We must now consider the more rapid changes which are observed in the angle of incidence of the downcoming waves. As an example of a case in which these variations were very marked, we may quote the following values, which were obtained by method (ii) within a period of 40 seconds on October 23, 1926, at 4.35 a.m.

$$\sin i_1 = 0{\cdot}60\,;\ 0{\cdot}48\,;\ 0{\cdot}50\,;\ 0{\cdot}46\,;\ 0{\cdot}46.$$

It appears reasonable to suppose that rapid variations such as these cannot be due to rapid variations in the height of the layer, since we do not observe corresponding variations in the number of interference fringes. On the contrary, it is found that the number of interference fringes remains remarkably constant, even when the angle of incidence may be varying widely. We may explain this most simply by supposing that "reflection" is taking place successively at different points on a layer, the mean height of which is approximately constant. This type of variation of angle of incidence might be expected if the layer were not of sensibly uniform horizontal stratification.

* 'Roy. Soc. Proc.,' A, vol. 113, p. 457 (1926).

We must not, however, exclude the possibility that the indirect ray not only arrives at the receiver in a downward direction, but is also laterally deviated. It is easily seen that such a lateral deviation would produce abnormally large values of sin i_1. If this lateral deviation occurs, we should expect it to be most marked at the transitional sunrise period, when we know that the layer must be tilted. Simultaneous experiments of the type described in this paper, carried out for transmissions in a north-south and in an east-west direction, during the sunrise period should yield information on this point.

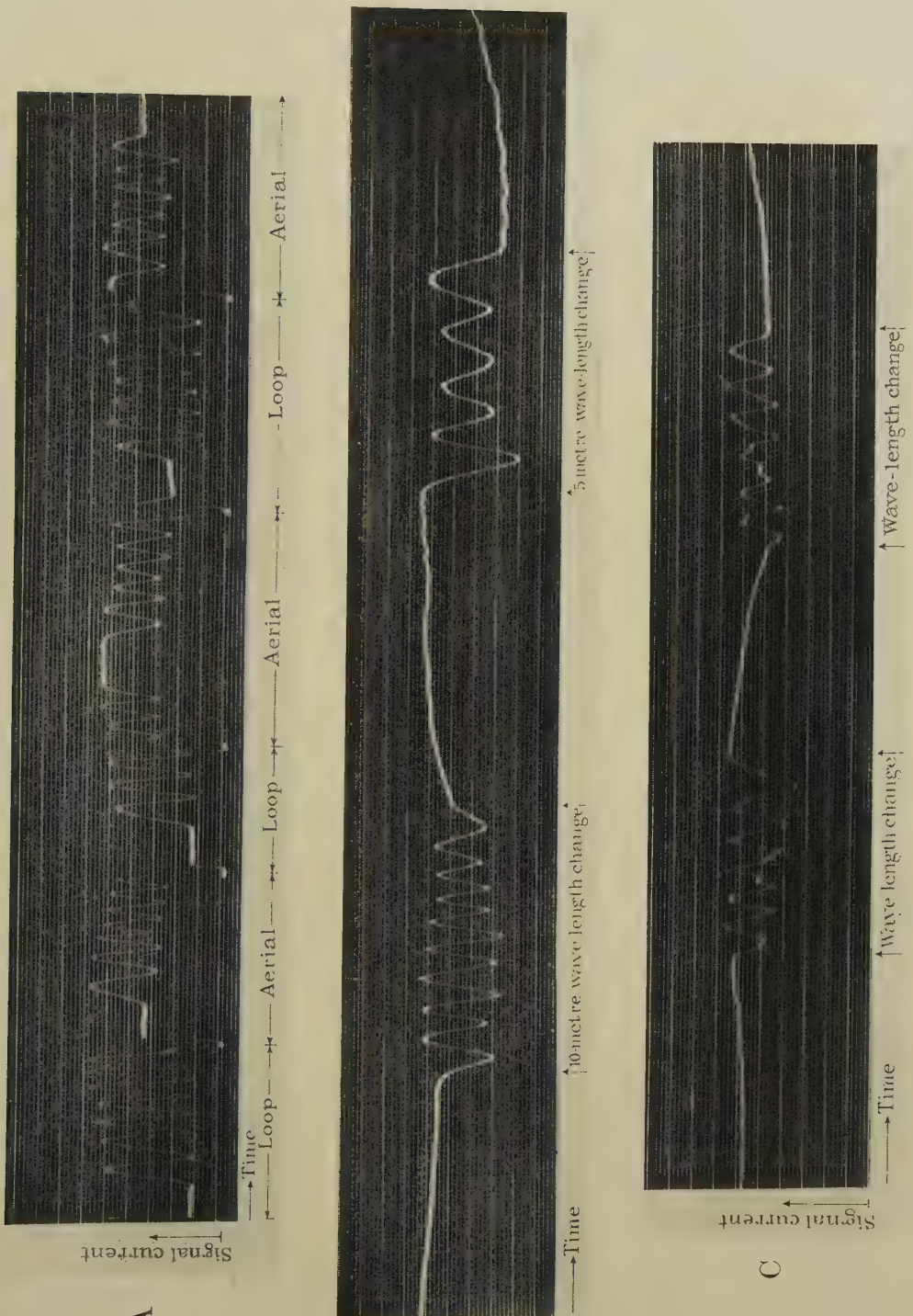
The fact that the angle of incidence is found to vary so rapidly, and over such a wide range, emphasizes a precaution which must be observed in determining sin i_1 by method (ii). It is essential that we should observe the fringes, produced on the loop and on the vertical aerial, in quick succession, so that sin i_1 does not alter appreciably between the two observations. A further reason for making the comparison as rapidly as possible is evident from Plate 5B, which shows that large and rapid changes may take place in the intensity of the downcoming wave. For example, the amplitude of the interference fringes (which is a measure of the indirect ray intensity), produced by the first wave-length change, is very different from that of the second set of fringes 10 seconds later. This is particularly evident if we compare the last fringe of the first wave-length change with the first fringe of the second. Since the rapid comparison of the signals on loop and vertical aerial is desirable, we have, in most cases, changed from one aerial to the other in the middle of a wave-length change (as in the third wave-length change of Plate 5A) rather than between two wave-length changes (as between the first and second changes of the same record).

Summary.

(1) Two methods of measuring the angle of incidence of downcoming wireless waves are described. The two methods have different ranges of applicability. Both involve photographic registration. The first method utilizes the ordinary night-time signal variations and can be employed in connexion with any steady transmitting station. It only yields useful results if the natural signal variations are small. The second method requires a controlled wave-length change at the transmitter, but may be used even when the natural signal variations are large.

(2) The mean values of the angle of incidence, as measured by these two methods, for the periods immediately following sunset and preceding sunrise, show a close agreement, and lead to an effective height of 90–100 kms. for the atmospheric deflecting layer.

(3) Observations of the angle of incidence, made by these methods, indicate


A
Signal current
Time
Loop
Aerial
Loop
Aerial
Loop
Aerial
B
Signal current
Time
10 metre wave length change
5 metre wave-length change
C
Signal current
Time
Wave length change
Wave-length change

(*Facing p.* 304.)

a diurnal variation in the height of the ionized layer, which is found to be higher in the middle of the night than during the sunset and sunrise periods.

(4) Comparatively rapid fluctuations have been observed in the angle of incidence of downcoming waves. Such fluctuations are not considered as being due to variations in the height of the ionized layer, but are explained by supposing that "reflection" takes place at different points on a layer the mean height of which is sensibly constant. Such variations might be expected if the layer were not of sensibly uniform horizontal stratification.

On the Nature of Wireless Signal Variations.—II.

By E. V. Appleton, M.A., D.Sc., F.R.S., Wheatstone Professor of Physics, King's College, London, and J. A. Ratcliffe, B.A., Clerk Maxwell Student, Cavendish Laboratory, Cambridge.

(Received May 30, 1927.)

1. *Introduction.*

The present paper continues the account of investigations on the nature of signal fading given in paper I of similar title. As the argument is essentially continuous from paper I to paper II the numbering of equations, tables and figures has been made to run continuously from the former to the latter paper.

2. *Direct Investigation of the Downcoming Waves by suppression of the Ground Waves.*

(*a*) *Theoretical Considerations.*—The ordinary curves of signal intensity variation obtained either with a vertical aerial or with a loop in the plane of propagation do not give us directly the variations in amplitude of the downcoming waves because we are unable to deduce from them whether a particular signal variation is due to a change of phase or a change of amplitude. We can, however, investigate the changes in intensity of the downcoming waves if we suppress the ground waves. This may be done by using the particular combination of loop and vertical aerial discussed below.

Let us consider first the case in which the ground waves alone are present and are received simultaneously on a vertical aerial and on a loop in the plane of propagation. Let the amplitudes of the electromotive forces introduced

into the two systems be $K_A E_0$ and $K_L E_0$ respectively. If a third circuit is coupled to these two systems it is possible to adjust the couplings and phase differences so that no resultant electromotive force is produced in it. To find the condition for such a balance we may, noting that the system is linear, imagine the two electromotive forces $K_A E_0 \varepsilon^{jpt}$ and $K_L E_0 \varepsilon^{jpt}$ introduced into the system separately. Suppose e_A is the electromotive force impressed on the third circuit due to the signal received on the aerial. Then

$$e_A = \frac{K_A E_0 \varepsilon^{jpt}}{\phi_A(D)}, \tag{11}$$

where D is the usual d/dt operator and its coefficients in ϕ_A are determined by the circuit constants. A similar relation,

$$e_L = \frac{K_L E_0 \varepsilon^{jpt}}{\phi_L(D)}, \tag{12}$$

holds for the electromotive force produced in the third circuit by the loop signal. If the effect on the third circuit is zero

$$e_A + e_L = \frac{K_A E_0 \varepsilon^{jpt}}{\phi_A(D)} + \frac{K_L E_0 \varepsilon^{jpt}}{\phi_L(D)} = 0. \tag{13}$$

At night, however, when downcoming waves are received the two electromotive forces (e_A and e_L) impressed on the third circuit are no longer equal and opposite. The resultant electromotive force impressed on this circuit is then e, where

$$e = \frac{K_A (E_0 \varepsilon^{jpt} + 2E_1 \sin i_1 \, \varepsilon^{j(pt+\theta)})}{\phi_A(D)} + \frac{K_L (E_0 \varepsilon^{jpt} + 2E_1 \varepsilon^{j(pt+\theta)})}{\phi_L(D)}. \tag{14}$$

But, substituting from (13), this becomes

$$e = \frac{2K_L E_1 (1 - \sin i_1) \, \varepsilon^{j(pt+\theta)}}{\phi_L(D)}. \tag{15}$$

Thus the amplitude of the electromotive force introduced into the third circuit at night is proportional to the intensity of the normally polarized component of the downcoming waves.

The discussion immediately above refers to the theoretical basis of a method of making a receiving system which responds only to the downcoming waves, but by a slight modification, which is most conveniently dealt with here, we can arrange that the system will receive only the ground waves. Let us suppose, as before, that the loop and vertical aerial are adjusted so as to introduce out-of-phase electromotive forces into the third circuit. Suppose also that we adjust the couplings so that

$$\frac{K_A \sin i_1}{\phi_A(D)} + \frac{K_L}{\phi_L(D)} = 0. \tag{16}$$

Then (14) by substitution from (16) becomes

$$e = \frac{K_A}{\phi_A(D)} E_0 (1 - \sin i_1) \epsilon^{jpt}. \quad (17)$$

Thus with such an adjustment the elimination of a downcoming ray, which arrives at the receiver with a constant angle of incidence, can be effected. The application of this method to the minimization of interference fading is considered below. It is also easily seen that we have here a method of measuring the angle of incidence of downcoming waves in a case in which the ground wave is absent (*e.g.* in short wave measurements).* Imagine such a case in which the coupling has been adjusted so that no signals are received. Then (16) holds. Suppose now that a separate experiment with a ground wave from a near-by station is made and, with the same adjustment, the ratio of e_L to e_A is measured. Then

$$\frac{e_L}{e_A} = \frac{K_L}{\phi_L(D)} \frac{\phi_A(D)}{K_A}. \quad (18)$$

Substituting from (16) we have

$$\frac{e_L}{e_A} = \sin i_1. \quad (19)$$

It would not be necessary to measure the ratio $\frac{e_L}{e_A}$ every time, for the loop circuit coupling with the third circuit could be calibrated in terms of this ratio once and for all.

(*b*) *Practical Arrangements.*—The practical arrangements used in the suppressed ground-ray system considered above are similar to those which have been described in connexion with direction-finding without bilateral ambiguity. The necessary phasing was produced either by suitably tuning the loop or vertical aerial as indicated in fig. 6, or by using the resistance phasing method of the Marconi Company (see fig. 7).† The latter method was usually employed. The vertical aerial and loop in the plane of propagation were used, and the apparatus was adjusted to give no signal in the day-time. A photographic record taken with this setting at night thus gave the variation of the normally-polarized downcoming wave. It was considered of interest to compare simultaneous records of the signals received on an ordinary vertical aerial and of the indirect ray received with the suppressed ground-ray system. In this case a short aperiodic aerial with a neutrodyne amplifier was used for

* Since this was written we have seen an advance copy of a paper by Mr. T. L. Eckersley, of the Marconi Company, in which the use of such a method is mentioned.

† *Vide* Keen, "Direction and Position Finding by Wireless," London, p. 49 *et seq.*

the former. Tests showed that the presence of this extra aerial did not disturb the balance of the suppressed ground ray system.

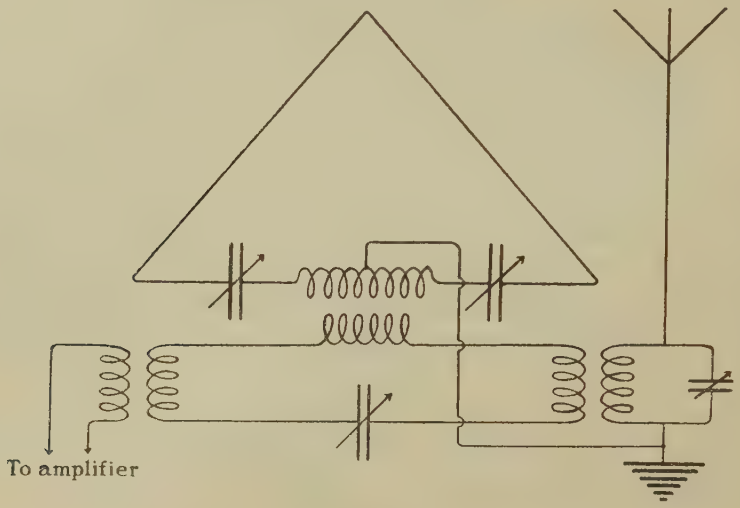


Fig. 6.

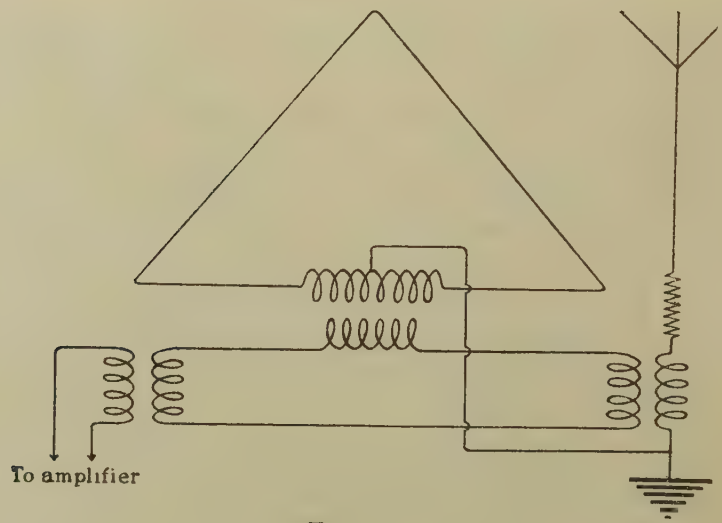
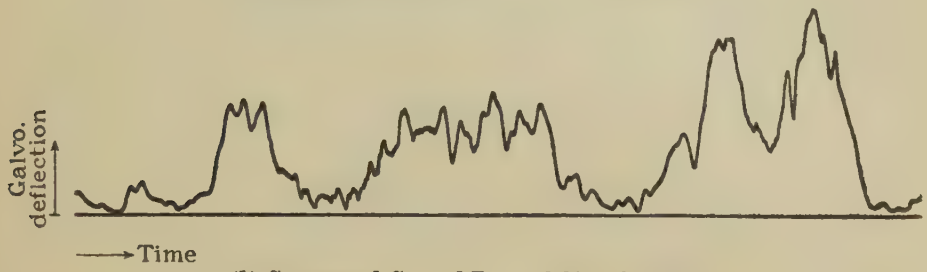


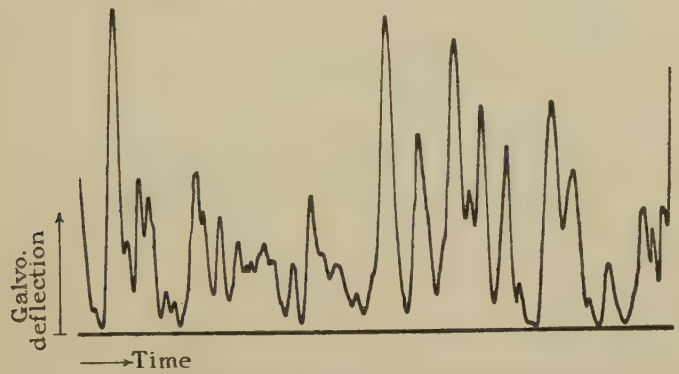
Fig. 7.

(*c*) *Experimental Results and Discussion.*—Fig. 8 shows a typical series of records obtained using the suppressed ground-ray system. They were taken on September 16, 1926, sunset being at 7.13 p.m. The four records (*a*), (*b*), (*c*) and (*d*) were taken at intervals between 9 p.m. and 11.0 p.m. In the case of record (*d*) the intensity of the indirect ray was so great that the overall

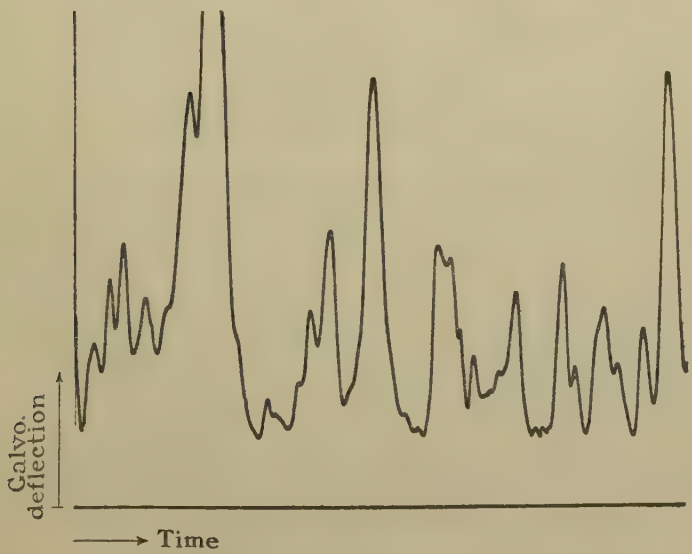
(*a*) Suppressed Ground Ray. 9.0 to 9.15 p.m.



(*b*) Suppressed Ground Ray. 9.30 to 9.45 p.m.

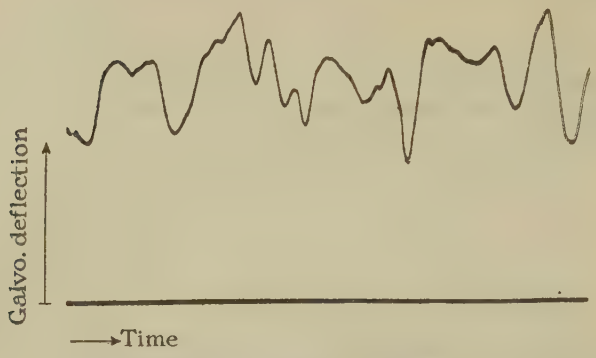


(*c*) Suppressed Ground Ray. 10.0 to 10.15 p.m.

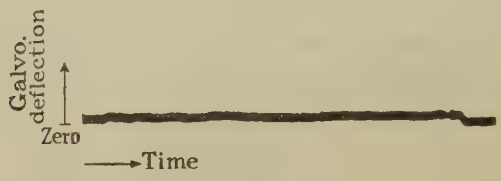


(*d*) Suppressed Ground Ray. 10.45 to 11.0 p.m.

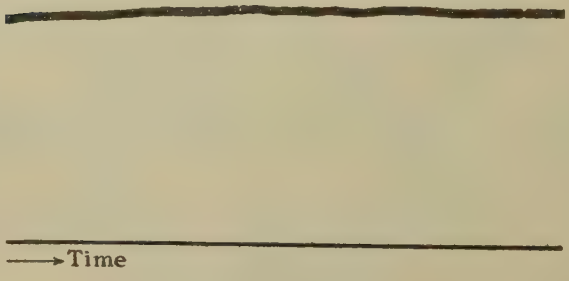
FIG. 8, (*a*) to (*d*).—Signal Currents (Suppressed Ground Ray System). 16.9.26.



(*e*) Vertical Antenna, beginning 10.55 p.m. 16.9.26.



(*f*) Suppressed Ground Ray. Midday, 17.9.26.



(*g*) Loop. Midday, 17.9.26.

FIG. 8, (*e*) to (*g*).—Signal Currents.

amplification of the receiver had to be reduced to keep the galvanometer deflection on the scale. Record (*e*) shows the variation in signal strength as received at 10.55 p.m. on a vertical aerial, and hence shows the kind of fading which is associated with the variations of the indirect ray shown in (*d*). Record (*f*) shows the signal obtained on the suppressed ground-ray system at mid-day the next day, with the same adjustment of apparatus and (*g*) shows the signal intensity on the loop antenna a few minutes afterwards.

This series of records illustrates the gradual increase of mean intensity of the downcoming ray as night progresses. About two hours after sunset the ray begins to be appreciable, and increases in intensity until about three and

a-half hours after sunset it has become comparable with the ground ray. A comparison of (*d*) and (*e*) shows that a strong downcoming wave is to be correlated with marked variation of normal signal intensity, while the records (*f*) and (*g*) taken the next day show that the absence of a downcoming ray is to be correlated with a steady signal.

It has been previously mentioned that fading may be due to changes in any one of the following variables which determine the nature of the downcoming waves.

(*a*) Angle of incidence.
(*b*) Intensity.
(*c*) Phase.
(*d*) Polarization.

We have already seen from the experiments described in paper I that large and rapid variations occur in the angle of incidence, and we might expect these changes to be responsible for some of the signal variations experienced on a vertical aerial. They cannot, however, be responsible for the fading on a loop aerial, since the signals received on a loop are independent of the angle of incidence of the downcoming wave. Thus, in a comparison of simultaneous signal records obtained on a vertical aerial and on a loop, we may say that any variation which occurs only on the former is due to a change in the angle of incidence. But an examination of such records as we have taken shows that every signal variation on the vertical aerial has its counterpart on the loop. This shows that no signal change is due simply to a variation of the angle of incidence. Since we have found that variations in the angle of incidence do occur we must suppose that such variations are masked by other changes (*e.g.*, in intensity) which accompany them.

To investigate the possible effect of changes in intensity and phase we have taken simultaneous records of the signals received on the suppressed ground-ray system and on a vertical aerial. Such simultaneous records are reproduced in fig. 9. As previously, the galvanometer deflections are measured in opposite directions.

From the figure it will be seen that, in most cases, changes in the vertical aerial signal correspond definitely to changes in the quantity $E_1 (1-\sin i_1)$ which is measured on the suppressed ground-ray system. Changes in the latter quantity may be due either to changes in E_1, the intensity of the downcoming waves, or to changes in $\sin i_1$, corresponding to a varying angle of incidence. We have already shown, however, that fading on the vertical aerial is not due

primarily to changes in the angle of incidence, and we must therefore conclude that such signal variations correspond to changes in the intensity of the down-

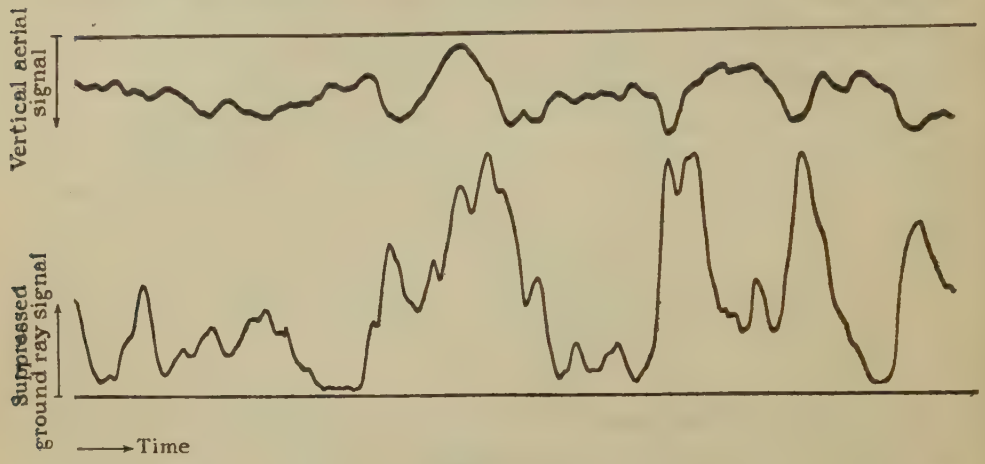


Fig. 9.

coming wave. Such fading may be called "intensity fading." Since we have seen that changes in sin i_1 are masked by the corresponding changes in intensity we shall, in what follows, speak of the signal received on the suppressed ground-ray system as indicating the intensity E_1 of the downcoming wave.

Bearing this in mind an examination of many records such as fig. 9 shows that an increase in E_1 is sometimes accompanied by an increase of signal intensity on the vertical aerial and sometimes by a decrease. This is explained by supposing that the relative phase of the direct and indirect waves is continually varying. When the two are in phase the resultant signal (received on the vertical aerial) increases when E_1 increases, but when they are out of phase the resultant signal decreases. Other evidence for changes in the relative phase of the direct and indirect waves is occasionally obtained on the "fringe" photograph. An example illustrating this is shown in Plate 5c, where the phase relation between the two waves is seen to alter very considerably during the pause (5 seconds) between the two wave-length changes. An examination of both types of records, however, shows that changes in the relative phase of ground and atmospheric waves ("phase fading") are less frequent than changes in the intensity of the indirect ray. It has previously been pointed out* that, when deducing E_1/E_0 and sin i_1 from the signal current curves obtained on loop and vertical aerial systems, we may not assume that each signal maximum is a phase maximum. The present results entirely justify this caution.

* Appleton and Barnett, 'Electrician,' *loc. cit.*

Although it has been shown that fading is mainly due to variations in the intensity of the normally polarized component E_1 of the downcoming waves it still remains to be determined whether such intensity changes are due to a wave of sensibly constant polarization, the amplitude of which is changing, or are to be ascribed to the varying state of polarization of a wave of constant amplitude. Experimental results bearing on this point have been obtained by taking simultaneous records of the signals received on the suppressed ground ray system and on a loop perpendicular to the direction of transmission. A sample set of such simultaneous records is shown in fig. 10.

As usual, the galvanometer deflections are recorded in opposite directions.

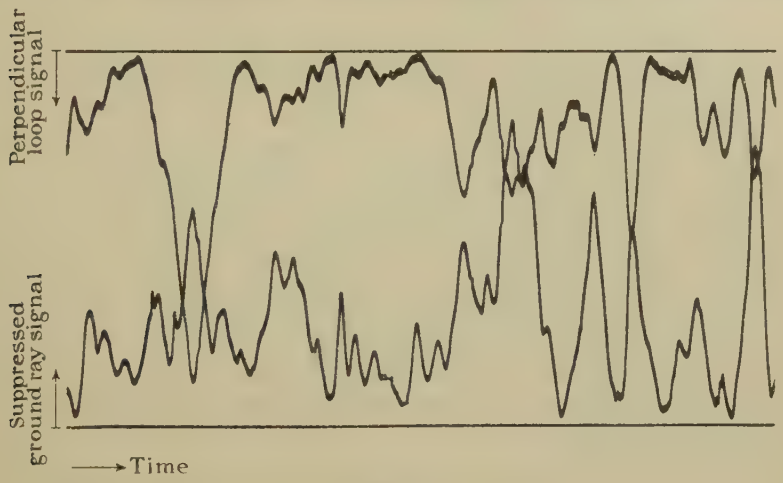


FIG. 10.

The upper record represents the signal variations received on the loop at right angles to the plane of propagation, and thus indicates the variation in $2H_1' \cos i_1$ (*i.e.* H_y). The lower record represents the variations of the signals received on the suppressed ground ray system, and thus indicates the variations of $E_1 (1 - \sin i_1)$. It will be seen that the signal current curves are remarkably similar as regards their main features. Now changes in i_1 would make the signals on the two systems increase or decrease together, and we must decide whether the similarity which is observed is mainly due to changes in i_1 or to the fact that E_1 and E_1' are varying together. We have seen above that there is strong evidence for supposing that the effect of any changes in i_1 are masked by the changes in E_1 which accompany them. We also find that on substituting the extreme values which we have found for i_1 into the expression $E_1' \cos i_1$ and $E_1 (1 - \sin i_1)$, that it is impossible to explain the large variations shown in

fig. 10 as due to changes in i_1. We therefore conclude that the two curves of this figure represent mainly the changes in intensity of the two components E_1 and E_1'. The fact that the two intensities increase and decrease together rules out the possibility of a plane polarized wave with a rotating plane of polarization, for in such a case an increase in the intensity of one component would be accompanied by a decrease in the intensity of the other. It should, however, be pointed out that the ratio of the signal variations on the two systems is by no means constant. This may be due either to a variation in the ratio E_1/E_1' or to a variation in the angle i_1. We cannot at present distinguish between these two possibilities.

We thus arrive at the following conclusions regarding the causes of signal variations observed on a vertical aerial at a distance of 80 miles from a transmitter of broadcasting wave-lengths. Fading is due to interference between a direct wave propagated along the ground and an indirect wave which is returned by ionic deflection in the upper atmosphere. The signal variations are due chiefly to variations in the intensity of the downcoming wave, and to a lesser degree to the variations in phase difference between the ground and atmospheric waves. The changes in the downcoming waves are real intensity changes (as measured at the ground), and are not due to the rotation of a plane polarized wave. Changes in the angle of incidence of the downcoming waves, although present, are not responsible in any marked degree for the signal variations, any effect which they may produce being masked by simultaneously occurring intensity changes.

The fact that the intensities of the normally and abnormally polarized components of the downcoming wave vary together seems to indicate that they are coherent, and are, in reality, components of a single wave rather than two waves which have been treated separately by the upper atmosphere. Such a result would be in accordance with the magneto-ionic theory, which would lead us to expect a downcoming ray which is, in general, elliptically polarized.*

Although the cause of signal variations has been traced mainly to intensity fluctuations of the downcoming rays as measured at the ground, we must not regard these fluctuations as being due to the alterations in reflecting power of a fixed flat surface, for other observations seem to indicate that these intensity

* Although the experiments described suffice to demonstrate that the downcoming ray possesses both normally and abnormally polarized components they give us no information relating to the phase difference between the two components. A separate series of experiments, which will form the subject of a future paper, have, however, shown that the phase difference $(\theta' - \theta)$ between the two components varies round about $\pi/2$ for transmissions from London to Peterborough.

fluctuations are themselves the result of an interference mechanism the nature of which is at present somewhat obscure. As an example of such observations the following may be quoted :—The intensity fluctuations themselves exhibit a kind of periodicity which seems to vary with the wave-length and with the distance of transmission. An empirical relation describing these phenomena has been found by Mr. A. L. Green (of King's College, London) and one of the writers. If T is the " period " of the fluctuations and λ the wave-length, T/λ is found to be a single valued function of the distance of transmission *d*. Such a relation strongly suggests that interference plays a part in causing these intensity variations. Various possible causes might be suggested, such as the simultaneous " reflection " of waves from two or more portions of a layer of non-uniform horizontal stratification, but it does not seem profitable to discuss the matter in greater detail at present.

3. *Experiments on the Elimination of Fading.*

The suppressed ground-ray system has been used to cut out the effect of the direct ray, and so enable us to study the indirect ray alone. It has been shown, however (see 2 (*a*)), that by a slight variation in the adjustments such a system may be arranged to eliminate the effects of a downcoming wave arriving at any angle. Since the indirect ray is responsible for nocturnal variations of signal strength, we might expect that, by means of such a system, we should be able to eliminate the cause of fading and receive a steady ground signal even at night-time.

Many attempts have been made to do this, and although a definite reduction in the percentage signal variation has been obtained, the system does not seem to be very promising as a method of eliminating fading completely. There is little doubt that this is largely due to the fact that with such a suppressed atmospheric ray system we are only able to eliminate a downcoming ray arriving at a constant angle, whereas in the actual case, as the other experiments have indicated, the angle of incidence of the downcoming ray is rapidly changing.

It is of interest to note that since such a suppressed atmospheric ray system will not receive waves incident on the ground at a certain angle, it is also incapable of radiating waves in the same direction. But the use of such a system for preventing the presence of an indirect ray at a given receiving station would most probably be frustrated by the occurrence of unsymmetrical ray paths mentioned above.

4. *The Relation between Signal Variations and Directional Errors.*

We now consider very briefly the bearing of these results on the cause of apparent directional variations. We have always found that the downcoming wave possesses normally and abnormally polarized components E_1 and E_1'. The presence of E_1' will cause directional errors, as was first pointed out by Bellini and Eckersley.

The fact that records of $E_1 (1 - \sin i_1)$ and $E_1' \cos i_1$ show similar variations establishes a correlation between fading, which is caused by E_1, and directional errors which are caused by E_1'. (The previous lack of such a correlation, which was due to the difficulty of unravelling the separate effects of ground and atmospheric waves, has often been used as an argument against the theory that both fading and directional errors are due to the Heaviside layer.) The question arises whether the abnormally polarized component E_1' is produced by the transmitter or is brought about by the deviation at the ionized layer. Now, the 2LO aerial system is a T-antenna with a horizontal portion situated approximately east and west, so that we should not expect the emission of an abnormally polarized wave in a north-south transmission direction. We therefore conclude that the abnormally polarized component is introduced by the reflection at the ionized layer. The results of Smith-Rose,* who found that directional errors were obtained even with a purely vertical transmitting aerial, may be interpreted in the same say.

The experiments described above were carried out as part of the programme of the Radio Research Board of the Department of Scientific and Industrial Research. We are deeply indebted to those who have co-operated in the carrying out of this series of experiments; to Mr. M. A. F. Barnett, for most valuable assistance at the receiving station; to Dr. R. L. Smith-Rose, Mr. E. L. Hatcher, and Mr. A. C. Haxton, who arranged and carried out the special transmissions from the National Physical Laboratory; to Mr. W. C. Brown, of the Peterborough Radio Research Station; and to the Chief Engineer of the British Broadcasting Corporation (Captain P. P. Eckersley) for his kind interest in these experiments and his good offices in arranging special early morning transmissions from 2LO.

Summary.

(1) An account is given of experiments designed to yield information on the nature of the variations of downcoming wireless waves, which are responsible for nocturnal signal variations. By employing a receiving assembly

* 'Journ. Inst. Elect. Eng.,' vol. 12, p. 957 (1924).

which is a combination of a loop and vertical aerial, it has been possible to eliminate the effects of the ground waves at the receiving station and to study the characteristics of the downcoming wave directly. Large variations in the intensity of the downcoming waves are found.

(2) It is pointed out that fading may be due to changes in any of the following variables which determine the nature of the downcoming waves.

(*a*) Angle of Incidence.
(*b*) Intensity.
(*c*) Phase.
(*d*) Polarization.

It is shown that for wave lengths of about 400 ms. and distances of about 80 miles, fading is chiefly due to changes in the intensity of the downcoming waves. Variations in the phase relation between ground and sky waves are a secondary cause of fading. Changes in the angle of incidence or polarization of the downcoming wave are not responsible in any very marked degree for signal variations.

(3) The downcoming ray has been shown to be of complex polarization, having electric vectors both in, and at right angles to, the plane of propagation. Similar intensity variations are found in both these vectors.

(4) The use of a suppressed atmospheric ray system in reception for the minimization of fading, and in transmission, for preventing the emission of upward rays, is discussed. Such a system may be used to find the angle of incidence of downcoming rays in the absence of direct rays.

The Dissociation of Carbon Dioxide at High Temperatures.

By R. W. Fenning and H. T. Tizard, C.B., F.R.S.

(Received May 4, 1927.)

The power of an internal combustion engine is greatest when operating with a "rich" mixture, that is to say, with a mixture which contains more fuel than is necessary for complete combustion. Similarly, it is found that if mixtures of carbon monoxide and air in varying proportions are exploded in a closed bomb at constant initial temperature and pressure, the explosion pressure is greatest when the ratio CO/O_2 is greater than 2. These phenomena are known to be connected with the dissociation of carbon dioxide at high temperatures, for if there were no dissociation we should expect the explosion pressure to be greatest when $CO/O_2 = 2$. No attention appears, however, to have been paid to the *position* of the maximum. It can be shown in the following way that there is a very simple relation between the composition of the mixture giving maximum pressure on explosion, and the dissociation of carbon dioxide at the maximum explosion temperature.

Let the initial composition be represented by the expression

$$2\,(1 + a)\,CO + O_2 + b\,N_2 \quad (\text{Total mols} = 3 + 2a + b),$$

and let P_i, T_i represent the initial pressure and temperature; P_e the maximum pressure observed after explosion, and T_e the corresponding maximum temperature.

The composition of the mixture at the maximum temperature may be represented thus:

$$2\,(1 - x)\,CO_2 + 2\,(a + x)\,CO + x\,O_2 + b\,N_2$$

$$(\text{Total mols} = 2 + 2a + x + b)$$

and

$$\frac{T_e}{T_i} = \frac{P_e}{P_i} \times \frac{3 + 2a + b}{2 + 2a + x + b}. \qquad (1)$$

Also, if equilibrium is reached we have

$$K_p = \left(\frac{P_{CO}}{P_{CO_2}}\right)^2 \times P_{O_2} = \left(\frac{a + x}{1 - x}\right)^2 \times \frac{x}{2 + 2a + x + b} \times P_e, \qquad (2)$$

where K_p is the dissociation constant at the temperature T_e, and P_{CO}, etc., the partial pressures of CO, etc.

Now if Q is the heat of combustion of carbon monoxide at the temperature T_i, the temperature reached on explosion is given by the equation

$$\frac{2Q\,(1-x)}{T_e - T_i} = 2\,(1-x)\,C_{CO_2} + 2\,(a+x)\,C_{CO} + xC_{O_2} + b\,C_{N_2},$$

where C_{CO_2}, etc., signify the mean specific heats at constant volume of CO_2, etc., between the temperatures T_e and T_i. If the mean specific heats of CO, O_2, N_2 are the same, this equation reduces to

$$\frac{2Q}{T_e - T_i} = 2C_{CO_2} + \frac{2a + 3x + b}{(1-x)}\,C_{N_2}. \tag{3}$$

When T_e is a maximum $dT_e/da = 0$. Hence

$$\frac{d}{da}\left(\frac{2a + 3x + b}{1-x}\right) = 0,$$

or

$$\frac{dx}{da} = -\frac{2\,(1-x)}{3 + 2a + b}. \tag{4}$$

Combining equations (1) and (2) we have

$$\left(\frac{a+x}{1-x}\right)^2 \times \frac{x}{3 + 2a + b} = \frac{K_p}{T_e} \times \frac{T_i}{P_i} = \text{a maximum when } T_e \text{ is a maximum, because } K_p \text{ increases more rapidly than } T_e.$$

Hence

$$\frac{d}{da}\left\{\left(\frac{a+x}{1-x}\right)^2 \times \frac{x}{3 + 2a + b}\right\} = 0. \tag{5}$$

By differentiating, and substituting the value for dx/da given by (4), we get finally

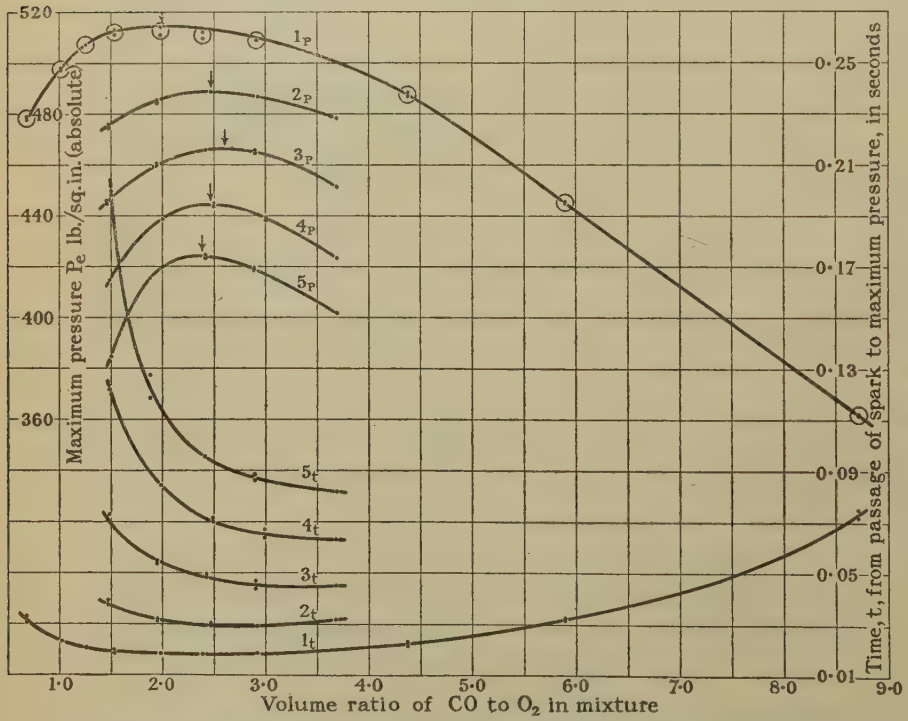
$$x = a/b \quad \text{when } T_e \textit{ is a maximum.} \tag{6}$$

It does not follow obviously that P_e is a maximum when T_e is a maximum, but this is easily shown to be the case by differentiating equation (1).

Hence $x = a/b$ when P_e is a maximum.

This conclusion is independent of any assumption as to the values of the mean specific heats of the gases. It rests on the assumption that the specific heats of the diatomic gases N_2, O_2, CO are the same at high temperatures. This is probably exactly true of CO and N_2, and it does not matter if the specific heat of O_2 is slightly different for the amount of O_2 present at the maximum temperature is small compared to that of CO and N_2. The conclusion is also independent of any assumption as to the amount of heat lost during the explosion. It assumes that there is no *variation* in the time of explosion of

mixtures on each side of the maximum point, which is not true in practice. The error introduced is, however, only significant when the ratio of N_2 to O_2 in the original mixture is high. Thus to determine the dissociation of carbon dioxide at high temperatures all that is necessary is to take a standard mixture of nitrogen and oxygen, to add carbon monoxide to it in varying quantities, and to determine the pressures reached on exploding the different mixtures. From the curve representing the results, the value of "a" when the explosion pressure is a maximum is read off; "x" is then given by equation (6), K_p by equation (2), and T_e by equation (1). By varying the proportion of nitrogen to oxygen in successive series of experiments it is possible to determine dissociation over a wide range of temperature. When no nitrogen is present, $b = 0$, and therefore $a = 0$ and x is indeterminate. This means that the pressure reached on exploding mixtures of carbon monoxide and oxygen is a maximum when $CO/O_2 = 2$. Experimental proof of this has been obtained (see fig. 1). It will be noted in this case that a wide variation in the proportion



Curve No.	$b = \frac{N_2}{O_2}$
1_P and 1_t	0·02 to
2_P „ 2_t	1·03
3_P „ 3_t	2·00
4_P „ 4_t	3·01
5_P „ 5_t	3·81

The arrows indi
the peaks of the cu

FIG. 1.—Maximum Pressures and Explosion Time for Mixtures of the approximate Composition $2(1 + a)\ CO + O_2 + b\ N_2 + 0{\cdot}01\ (3 + 2a + b)\ H_2O$.

Note.—The small H_2 content has been included in the CO.

of CO to O_2 has very little effect on the magnitude of the explosion pressure, indicating qualitatively that dissociation is considerable.

The Tables at the end of the paper contain the results of a number of experiments* that have been carried out. From the corresponding curves in fig. 1 the following results are deduced :—

Table I.

b.	a.	$x = a/b$.	P_e.		T_e from equation (1).	K_p from equation (2).
			lbs./sq. in. (absolute).	Atmos.		
1·03	0·235	0·23	488	33·2	3320	0·75
2·0	0·30	0·15	466	31·7	3095	0·28
3·01	0·23	0·076	444	30·2	2920	0·046
3·81	0·19	0·05	423	28·8	2750	0·015

The chief error in determining K_p by this method is involved in the difficulty of judging exactly the point of maximum pressure. A 10 per cent. increase in "a" corresponds to a 35 per cent. increase in K_p, and a *decrease* in the associated maximum temperature of less than 1 per cent. At these high temperatures K_p increases by 20 to 30 per cent. for every 1 per cent. rise in absolute temperature, so the cumulative effect of a 10 per cent. error in "a" is to produce an error in the determination of dissociation by this method which corresponds to a 2 per cent. error in temperature, and is therefore less important than it appears at first sight. The experimental results were so consistent throughout that it is not considered likely that the error due to this cause in determining "a" is greater than 10 per cent.

Another error is introduced by the circumstance that the time of explosion, and therefore the loss of heat during explosion, increases as the proportion of carbon monoxide decreases. The apparent value of "a" at the point of maximum pressure is therefore greater than the true value, and the dissociation

* The experiments were carried out at the N.P.L. for the Engineering Research Board of the Department of Scientific and Industrial Research. The apparatus used has already been described by one of us (R.W.F.) in 'Phil. Trans.,' A, vol. 225, pp. 331–356, wherein are also given the results of a series of CO-air and H_2-air explosions. A repetition of some of the latter experiments with the present setting of the manometer resulted in a 2 per cent. increase in the maximum pressures previously recorded—and this discrepancy must be borne in mind when making any comparisons between the two series of experiments. The higher value is supported by still earlier work. The cause of the discrepancy has not been determined.

constants given in Table 1 tend to be *too high*. This error is negligible at the highest temperature ($b = 1{\cdot}03$), but quite significant at the lowest ($b = 3{\cdot}8$). The results have therefore been checked by another method which, in principle, has been used by Bjerrum and subsequent investigators. If two mixtures containing different proportions of CO, O_2 and N_2 give the same pressure on explosion, then, on the same reasoning as before, they must contain the same proportion of carbon dioxide at the maximum temperature, if the loss of heat during explosion is the same in the two cases. From this we can calculate K_p in a way best shown by an example.

The mixture containing* (*vide* Table VIII, Record No. 740A)

$$0{\cdot}3825\ CO + 0{\cdot}1535\ O_2 + 0{\cdot}462\ N_2 + 0{\cdot}002\ CO_2 + 0{\cdot}011\ H_2O$$

(Total mols. 1·011)

gave the same pressure on explosion as the mixture (*vide* Table V, Record No. 671)

$$0{\cdot}8345\ CO + 0{\cdot}1415\ O_2 + 0{\cdot}0195\ N_2 + 0{\cdot}0045\ CO_2 + 0{\cdot}011\ H_2O$$

(Total 1·011).

In view of the large excess of carbon monoxide in this latter mixture, it can be assumed in the first place that all the oxygen is burnt, and therefore that 0·283 mols of CO_2 are formed. Hence, after explosion of the first mixture we should have

$$0{\cdot}283\ \text{mols}\ CO_2 + 0{\cdot}0995\ CO + 0{\cdot}012\ O_2 + 0{\cdot}462\ N_2 + 0{\cdot}002\ CO_2 + 0{\cdot}011\ H_2O$$

(Total mols. 0·870),

and

$$K_p = \left(\frac{0{\cdot}0995}{0{\cdot}285}\right)^2 \times \frac{0{\cdot}012}{0{\cdot}870} \times P_e$$

$$= 0{\cdot}051 \qquad (P_e = 30{\cdot}2\ \text{atmospheres}),$$

also

$$T_e = \frac{P_e}{P_i} \times T_i \times \frac{1{\cdot}011}{0{\cdot}87} = 2910^\circ\ \text{absolute}.$$

If the calculation is carried out more accurately, by allowing for the slight dissociation in the mixture containing excess of carbon monoxide, we get

$$K_p = 0{\cdot}054.$$

This is an upper limit for the value of K_p at this temperature, since the difference in loss of heat during the explosion of the two mixtures is ignored. The explosion times (from commencement of rise of pressure) were 0·027 seconds

* The small amount of H_2 present in the mixture has been counted as CO for the purpose of this calculation. The error introduced by this assumption is negligible.

and 0·060 seconds respectively. From a comparison of the cooling curves, and from other data, it is estimated that the loss of heat was greater in the second case by an amount of the order of 1 per cent., and probably not more than $1\frac{1}{2}$ per cent. of the total heat evolved. Hence to get a *lower* limit for the dissociation constant it is necessary to compare the first mixture with one containing excess of CO and giving a $1\frac{1}{2}$ per cent. greater pressure on explosion. Alternatively we can compare the same CO/O_2 mixture with the mixture numbered 741 in Table VIII, which gave a maximum pressure $1\frac{1}{2}$ per cent. lower. If we do this we calculate that

$$K_p = 0{\cdot}025 \text{ at } T_e = 2870.$$

Carrying out similar calculations by comparing various mixtures we have obtained the results given in Table II.

Table II.

Mixtures compared.	P_e.		T_e C° (abs.).	K_p.	
	lbs./sq. in. (absolute).	Atmos.		Probable upper limit.	Probable lower limit.
727 and 674	488	33·2	3310	0·69	—
729 and 674	484	33·0	3295	—	0·48
735 and corresponding CO/O_2 mixture	465	31·7	3100	0·20	—
735 and CO/O_2 mixture, giving 1 per cent. greater P_e	465	31·7	3115	—	0·14
740A and 671	444	30·2	2910	0·054	—
741 and 671	438	29·8	2870	—	0·025
744 and CO/O_2 mixture, giving 1 per cent. greater P_e	424	28·9	2745	0·018	—
744 and CO/O_2 mixture, giving 3 per cent. greater P_e	424	28·9	2760	—	0·0035

It will be seen by comparing Tables I and II that the dissociation constants calculated by the two methods agree very closely. One would expect the variation of K_p with temperature over this range to be expressed by an empirical equation of the form

$$\log K_p = A - \frac{B}{T}.$$

By trial it is found that the equation

$$\log K_p = 8{\cdot}46 - \frac{28600}{T}, \qquad (7)$$

not only reproduces accurately the experimental results communicated in this paper, but also those obtained at lower temperatures by other methods. This is shown in Table III. In the last two columns of this table we compare the degree of dissociation "y" of CO_2 at atmospheric pressure, calculated from the equation

$$\left(\frac{y}{1-y}\right)^2 \times \frac{y}{2+y} = K_p,$$

with those measured directly at lower temperatures or calculated from the values of K_p deduced from our explosion experiments.

Table III.

T/C° (abs.).	K_p calculated from equation (7).	K_p found.	100 y from K_p "calculated."	100 y found.
3320	0·70	0·75	63	63·5
3310	0·66	<0·69	62·5	<63
3295	0·60	>0·48	61·5	>59
3115	0·19	>0·14	49·5	>46·5
3100	0·17	<0·20	48·5	<50
3095	0·166	0·28	48	53·5(?)
2920	0·046	0·046	35·5	35·5
2910	0·043	<0·054	35	<37
2870	0·031	>0·025	32	>30
2760	0·013	>0·0035	25·5	>17
2750	0·0115	0·015	24·5	26·5
2745	0·011	<0·018	24	<28
2420	$4·4 \times 10^{-4}$	—	9·1	10 to 11*
2240	$4·9 \times 10^{-5}$	—	4·5	4·5*
1820	$5·6 \times 10^{-8}$	—	0·48	0·45†
1565	$1·5 \times 10^{-10}$	—	0·067	0·064‡
1500	$2·5 \times 10^{-11}$	—	0·037	0·047‡(?)
1480	$1·4 \times 10^{-11}$	—	0·030	0·028‡ 0·029 to 0·035§
1440	$4·0 \times 10^{-12}$	—	0·020	0·025‡
1400	$1·1 \times 10^{-12}$	—	0·013	0·01 to 0·02§
1300	$2·9 \times 10^{-14}$	—	0·0039	0·0041§

* Emich, 'Monatshefte f. Chem.,' vol. 26, p. 1011 (1905).
† Löwenstein, 'Z. f. physik. Chem.,' vol. 54, p. 711 (1906).
‡ Langmuir, 'Journ. Amer. Chem. Soc.,' vol. 28, p. 1379 (1906).
§ Nernst and von Wartenberg, 'Z. f. physik. Chem.,' vol. 56, p. 548 (1906).

Our results for the dissociation of carbon dioxide at high temperatures are considerably lower than those of Bjerrum.* For instance, according to Bjerrum, carbon dioxide is half dissociated at atmospheric pressure ($K_p = 0·2$) at a temperature of 2940° C. absolute, whereas our corresponding temperature is 3130, or about 200 deg. higher. At 3300 deg. absolute, K_p according to Bjerrum is 2·5, from which $y = 75$ per cent.; whereas we find $K_p = 0·62$ and

* 'Z. f. physik. Chem.,' vol. 79, p. 537 (1912).

$y = 62$ per cent. Nor do we agree with Crowe and Newey who found that a mixture of 40 per cent. CO with 60 per cent. air gave the maximum rise of pressure on explosion, whereas we find the proportion of CO to be 33 per cent. We think that these differences can be fully accounted for by the fact that previous investigators have found it extremely difficult to obtain accurate measurements of explosion pressures. Crowe and Newey state that their measurements, on any single mixture, differed by as much as 2 per cent. from the mean value. Bjerrum refers to the variation of the times of explosion of mixtures in successive experiments and attributes the cause rightly to the great influence of water vapour. It is indeed impossible to deduce dissociation constants from explosion experiments with any confidence unless the results of successive experiments with any mixture are very consistent. To secure consistency it is absolutely necessary to keep the moisture content constant. One example will suffice to show the great influence of water vapour.

Table IV.

Record No.	Composition of Mixture.						P_e.	Time intervals—seconds.	
	CO.	O_2.	N_2.	CO_2.	H_2.			Spark to maximum.	Commencement of rise to maximum.
664	87·35	10·05	1·90	0·45	0·25		356	0·1506	0·124
665						+ 0·02 H_2O	355	0·1535	0·128
666						+ 1·06 H_2O	361	0·0751	0·065
667							362	0·0720	0·063

It will be observed that the addition of over 1 per cent.* of water vapour lowered the time of explosion by over 0·06 seconds and raised the explosion pressure by $1\frac{1}{2}$ per cent. These results lead to a useful estimation of the loss of heat during explosion. We should expect the addition of 1 per cent. water vapour to *lower* the explosion pressure by over 1 per cent. if there were no loss of heat during the explosion, and no variation in dissociation or the completeness of combustion. The fact that it raises it by $1\frac{1}{2}$ per cent. means that the difference in the explosion times of 0·06 seconds corresponds to a difference in heat loss of about 3 per cent. of the total heat evolved. It is on this assump-

* From experiment it appears that about 30 per cent. of the added water vapour is absorbed by the explosion vessel walls (in their present condition), and will not therefore take part in the combustion until the flame reaches those surfaces.

tion, and on the further assumption that the difference in heat losses is proportional to the difference in times of explosion, that the lower limits for K_p (Table II) have been calculated. Check calculations, based on a somewhat different method, and taking into account the variation in dissociation due to the water vapour addition to the dry mixture, gave substantially similar results. Estimates of heat losses from the cooling curves after the maximum pressure has been reached lead to rather lower values.

It will be seen from the data given in Tables V to X how closely maximum pressures and times of explosion agree in successive experiments on the same mixtures. Several pressure-time records have been reproduced (in part) in fig. 2; of these Nos. 685 and 687 are for the same mixture, except that in the latter there is practically no water vapour.

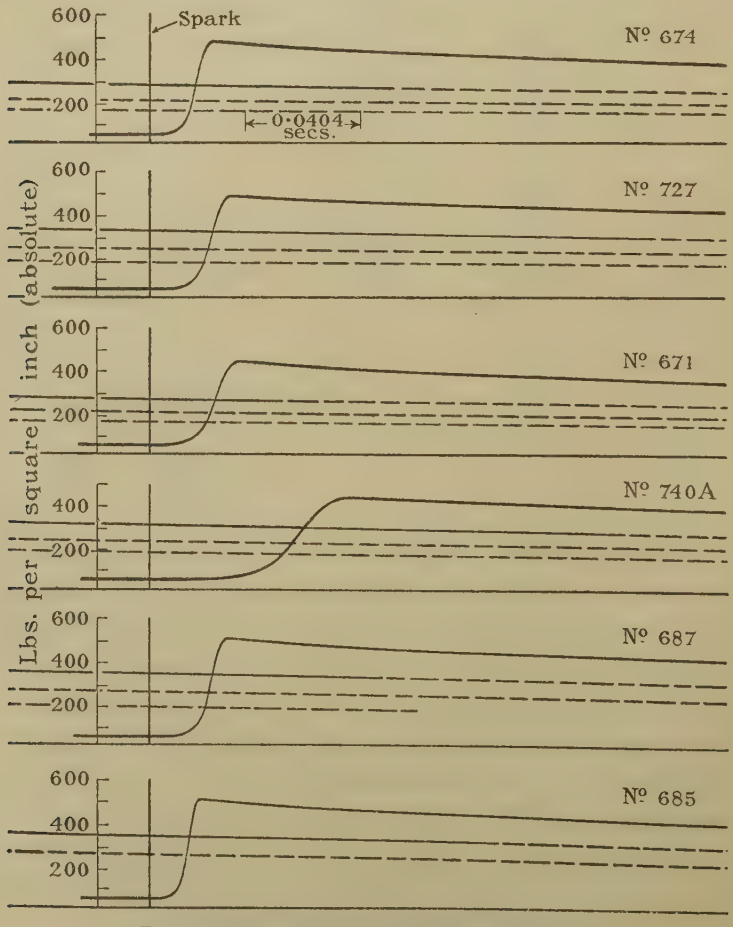


Fig. 2.—Pressure—Time—Explosion Records.

Table V.

Mixture $2(1 + a)\,CO + O_2$. Initial Pressure (abs.) 57·3 lbs./sq. in. (3·9 atmos.). Initial Temp. 50° C.

Record No.	Mixture. Composition per cent. by volume. CO.	O_2.	N_2.	CO_2.	H_2.	Ratio $\frac{CO^*}{O_2}$.	Volumes of water vapour added† per 100 vols. of mixture.	Products of Combustion. Partial analysis per cent. by volume. CO_2.	O_2.	CO.	Maximum Explosion Pressure (absolute). lbs. per sq. in.	Atmos.	Pressure 1/10 sec. after maximum (abs.) lbs. per sq. in.	Time Intervals. Spark to Maximum Pressure seconds.	Commencement of rise to Maximum Pressure seconds.
666	87·35	10·05	1·90	0·45	0·25	8·71	1·06	—	—	—	361	24·6	325	0·0750	0·065
667							1·06	—	—	—	362	24·65	327	0·0720	0·063
670	83·20	14·15	1·95	0·45	0·25	5·91	1·06	32·95	0·15	64·1	445	30·3	380	0·0323	0·027
671							1·08	—	—	—	444	30·25	380	0·0318	0·027
674	79·05	18·15	2·0	0·55	0·25	4·37	1·06	—	—	—	488	33·2	417	0·0225	0·019
675							1·06	44·3	0·1	53·1	487	33·15	417	0·0216	0·018
678	72·55	24·95	1·75	0·45	0·3	2·92	1·05	66·2	0·15	31·1	508	34·6	444	0·0184	0·016
679							1·06	—	—	—	508	34·6	444	0·0182	0·016
682	68·70	28·85	1·75	0·45	0·3	2·39	1·06	80·85	0·15	16·6	511	34·85	446	0·0182	0·014
683							1·07	—	—	—	509	34·7	446	0·0182	0·014
684	64·8	32·8	1·7	0·4	0·3	1·99	1·08	97·65	0	0	510	34·75	445	0·0180	0·015
685							1·06	—	—	—	513	35·0	447	0·0180	0·015
696	59·0	38·65	1·65	0·35	0·35	1·53	1·06	85·75	12·1	0·25	511	34·8	447	0·0184	0·015
697							1·06	—	—	—	512	34·9	446	0·0192	0·016
700	54·45	43·4	1·5	0·3	0·35	1·26	1·04	76·45	21·5	0·35	507	34·5	445	0·0203	0·016
701							1·06	—	—	—	507	34·5	445	0·0207	0·016
704	49·3	48·6	1·45	0·3	0·35	1·02	1·06	66·9	31·3	0·3	497	33·85	437	0·0230	0·018
705							1·06	—	—	—	497	33·85	437	0·0227	0·018
721	39·85	58·05	1·45	0·25	0·4	0·69	1·07	50·95	47·35	0·4	478	32·55	424	0·0302	0·024
722							1·07	—	—	—	477	32·5	424	0·0322	0·027

* Including H_2.

† Including an allowance for water vapour in orignal mixture. *Vide* Footnote to page 325 for absorption.

Table VI.

Mixture 2 $(1 + a)$ CO + O_2 + 1·03 N_2. Initial Pressure (abs.) 57·3 lbs./sq. in. (3·9 atmos.). Initial Temp., 50° C.

Record No.	Mixture. Composition per cent. by volume. CO.	O_2.	N_2.	CO_2.	H_2.	Ratio $\frac{CO^*}{O_2}$.	Volumes of water vapour added† per 100 vols. of mixture.	Products of Combustion. Partial analysis per cent. by volume. CO_2.	O_2.	CO.	Maximum Explosion Pressure (absolute). lbs. per sq. in.	Atmos.	Pressure 1/10 sec. after maximum (abs.) lbs. per sq. in.	Time Intervals. Spark to Maximum Pressure seconds.	Commencement of rise to Maximum Pressure seconds.
725	58·65	20·05	20·7	0·4	0·2	2·93	1·1	49·85	0·15	23·9	486	33·1	431	0·0285	0·023
726							1·1	—	—	—	486	33·1	432	0·0289	0·024
727	54·45	22·10	22·85	0·4	0·2	2·47	1·1	56·65	0·05	13·8	488	33·25	436	0·0297	0·023
728							1·1	—	—	—	488	33·25	437	0·0291	0·023
729	48·55	25·0	25·85	0·4	0·2	1·95	1·1	65·35	0·25	0	484	33·0	435	0·0307	0·025
730							1·1	—	—	—	485	33·0	435	0·0315	0·025
731	41·8	28·3	29·35	0·35	0·2	1·48	1·1	53·95	8·55	0	474	32·3	429	0·0372	0·031
732							1·1	—	—	—	475	32·35	429	0·0389	0·033

* Including H_2.

† Including an allowance for water vapour in original mixture. *Vide* footnote to page 325 for absorption.

Table VII.

Mixture 2 $(1 + a)$ CO + O_2 + $2N_2$. Initial Pressure (abs.) 57·3 lbs./sq. in. (3·9 atmos.). Initial Temp., 50° C.

Record No.	Mixture. Composition per cent. by volume. CO.	O_2.	N_2.	CO_2.	H_2.	Ratio $\frac{CO^*}{O_2}$.	Volumes of water vapour added† per 100 vols. of mixture.	Products of Combustion. Partial analysis per cent. by volume. CO_2.	O_2.	CO.	Maximum Explosion Pressure (absolute). lbs. per sq. in.	Atmos.	Pressures 1/10 sec. after maximum (abs.) lbs. per sq. in.	Time Intervals. Spark to Maximum Pressure seconds.	Commencement of rise to Maximum Pressure seconds.
733	48·85	16·9	33·75	0·35	0·15	2·90	1·1	40·55	0·1	18·65	464	31·6	415	0·0434	0·035
734							1·1	—	—	—	465	31·65	415	0·0462	0·038
735	44·5	18·4	36·6	0·3	0·15	2·42	1·1	44·9	0·1	10·0	465	31·7	421	0·0480	0·039
736							1·1	—	—	—	465	31·7	420	0·0487	0·039
737A	39·1	20·2	40·3	0·25	0·15	1·94	1·1	49·2	0·35	0	459	31·3	418	0·0534	0·043
737B							1·1	—	—	—	460	31·3	418	0·0543	0·041
738A	32·5	22·4	44·7	0·25	0·15	1·46	1·1	39·65	6·5	0	444	30·2	403	0·0696	0·061
738B							1·1	—	—	—	445	30·3	404	0·0722	0·062

* Including H_2.

† Including an allowance for water vapour in original mixture. *Vide* footnote to page 325 for absorption.

Table VIII.

Mixture 2 $(1 + a)$ CO + O_2 + 3·01 N_2. Initial Pressure (abs.) 57·3 lbs./sq. in. (3·9 atmos.). Initial Temp., 50° C.

Record No.	Mixture. Composition per cent. by volume. CO.	O_2.	N_2.	CO_2.	H_2.	Ratio $\frac{CO^*}{O_2}$.	Volumes of water vapour added† per 100 vols. of mixture.	Products of Combustion. Partial analysis per cent. by volume. CO_2.	O_2.	CO.	Maximum Explosion Pressure (absolute). lbs. per sq. in.	Atmos.	Pressure 1/10 sec. after Maximum (abs.) lbs. per sq. in.	Time Intervals. Spark to Maximum Pressure seconds.	Commencement of rise to Maximum Pressure. seconds.
739A	42·5	14·25	42·9	0·25	0·1	2·99	1·1	33·25	0·1	16·6	439	29·85	392	0·0635	0·054
739B							1·1	—	—	—	438	29·85	392	0·0665	0·057
740A	38·15	15·35	46·2	0·2	0·1	2·49	1·1	36·3	0·05	9·0	444	30·25	402	0·0707	0·060
740B							1·1	—	—	—	443	30·2	402	0·0694	0·060
741A	32·95	16·65	50·05	0·25	0·1	1·99	1·1	39·65	0·05	0·2	438	29·8	401	0·0839	0·069
741B							1·1	—	—	—	438	29·8	401	0·0837	0·073
742A	27·0	18·2	54·55	0·15	0·1	1·49	1·1	32·0	4·5	0	414	28·15	380	0·1210	0·103
742B							1·1	—	—	—	414	28·2	381	0·1211	0·101

* Including H_2.

† Including an allowance for water vapour in original mixture. *Vide* footnote to page 325 for absorption.

Table IX.

Mixture 2 $(1 + a)$ CO + O_2 + 3·81 N_2. Initial Pressure (abs.) 57·3 lbs./sq. in. (3·9 atmos.). Initial Temp. 50° C.

Record No.	Mixture. Composition per cent. by volume. CO.	O_2.	N_2.	CO_2.	H_2.	Ratio $\frac{CO.*}{O_2}$.	Volumes of water vapour added† per 100 vols. of mixture.	Products of Combustion. Partial analysis per cent. by volume. CO_2.	O_2.	CO.	Maximum Explosion Pressure (absolute). lbs. per sq. in.	Atmos.	Pressure 1/10 sec. after maximum (abs.) lbs. per sq. in.	Time Intervals. Spark to Maximum Pressure seconds.	Commencement of rise to Maximum Pressure seconds.
743A	37·35	12·95	49·35	0·25	0·1	2·89	1·1	29·7	0·1	13·55	418	28·5	377	0·0862	0·072
743B							1·1	—	—	—	419	28·55	378	0·0877	0·074
744A	33·25	13·8	52·65	0·25	0·05	2·41	1·1	31·95	0·05	6·8	423	28·8	385	0·0946	0·080
744B							1·1	—	—	—	424	28·9	387	0·0947	0·081
745A	28·15	14·95	56·65	0·2	0·05	1·88	1·15	33·3	0·5	0	415	28·3	383	0·1180	0·099
745B							1·1	—	—	—	415	28·25	384	0·1269	0·109
748	23·7	15·8	60·25	0·2	0·05	1·51	1·1	27·55	3·65	0	385	26·2	356	0·1980	0·167
749							1·1	—	—	—	384	26·2	357	0·1991	0·173

* Including H_2.

† Including an allowance for water vapour in original mixture. *Vide* footnote to page 325 for absorption.

Table X.

Mixture $3{\cdot}69\,CO + O_2 + bN_2$. Initial Pressure (abs.) 57·3 lbs./sq. in. (3·9 atmos.). Initial Temp., 50° C.

Record No.	Mixture. Composition per cent. by volume. CO.	O_2.	N_2.	CO_2.	H_2.	Ratio $\frac{CO^*}{O_2}$.	Volumes of water vapour added† per 100 vols. of mixture.	Products of Combustion. Partial analysis per cent. by volume. CO_2.	O_2.	CO.	Maximum Explosion Pressure (absolute). lbs. per sq. in.	Atmos.	Pressure 1/10 sec. after Maximum (abs.) lbs. per sq. in.	Time Intervals. Spark to Maximum Pressure seconds.	Commencement of rise to Maximum Pressure seconds.
750	63·9	17·4	18·0	0·5	0·2	3·68	1·1	42·1	0·05	35·9	478	32·6	418	0·0313	0·026
751							1·1	—	—	—	479	32·65	419	0·0333	0·028
752	54·6	14·9	29·7	0·6	0·2	3·68	1·2	35·05	0·1	29·85	451	30·7	394	0·0437	0·036
753							1·2	—	—	—	451	30·75	396	0·0450	0·037
754	45·15	12·3	41·8	0·6	0·15	3·69	1·2	—	—	—	412	28·05	367	0·0722	0·060
755							1·2	28·2	0·1	23·95	413	28·1	367	0·0724	0·061
756	42·85	11·6	44·75	0·65	0·15	3·70	1·2	26·65	0·1	22·55	398	27·1	357	0·0806	0·068
757							1·2	—	—	—	400	27·2	358	0·0829	0·069

* Including H_2.

† Including an allowance for water vapour in original mixture. *Vide* footnote to page 325 for absorption.

In conclusion, it should be remarked that the relation $x = a/b$ at the maximum temperature applies also to flames. Thus, by burning carbon monoxide or hydrogen in a Bunsen burner, and determining the composition of the mixture which gives the maximum flame temperature, it should be possible to make direct measurements of the dissociation of water vapour and carbon dioxide at high temperatures and at atmospheric pressure.

Summary.

It has been shown that the dissociation of carbon dioxide at high temperatures can be deduced very simply from measurements of the composition of mixtures of carbon monoxide, oxygen and nitrogen which give the greatest rise in pressure on explosion in a closed vessel, when the initial conditions of pressure and temperature are kept constant. The dissociation constants found from our explosion experiments by this method agree closely with those deduced from the same experiments by the method used hitherto.

The results obtained for the dissociation of carbon dioxide at high temperatures are considerably lower than the accepted values.

The authors' thanks are due to Mr. F. T. Cotton (N.P.L.) for his valuable assistance.

The Equation of State of a Gaseous Mixture.

By J. E. Lennard-Jones, D.Sc., Reader in Mathematical Physics, and W. R. Cook, B.Sc., Research Student, The University, Bristol.

(Communicated by R. H. Fowler, F.R.S.—Received March 15, 1927.)

§ 1. *Introduction.*

The methods of determining intermolecular fields, which have been used in former papers,* apply only to molecules of the same kind. This paper considers the more general problem of determining the forces between molecules of different kinds. It shows that this information can be derived from a study of the equation of state of binary mixtures, provided that the experimental work covers a sufficiently wide range of temperature and deals with mixtures of sufficiently varied proportions. The experimental information available at present is somewhat scanty and needs considerable extension.

Apart from the earlier measurements of Amagat, the only available work on gaseous mixtures appears to be that of Holborn and Olto† on a mixture of helium and neon, that of Verschoyle‡ on a mixture of hydrogen and nitrogen, and that of Masson and Dolley§ on binary mixtures of ethylene, argon and oxygen. The former carried out experiments on a mixture of one proportion only, but have supplemented the results with further work on helium alone and neon alone, in each case covering the range of temperature 0 to 400° C. This work provides just enough data for the purpose of this paper, and from it the field between a helium atom and a neon atom, as well as the fields between two helium atoms and two neon atoms, can be deduced.

Verschoyle, on the other hand, dealing with hydrogen and nitrogen, has worked at two temperatures only, viz., 0 and 20°, but has extended his observations in another direction. He has examined the successive changes in the isothermal of the mixture as its proportions are gradually changed. These changes are considered in a quantitative way in this paper and are satisfactorily

* ' Roy. Soc. Proc.,' A, vol. 106, pp. 441 and 463 (1924); vol. 107, p. 157 (1925); vol. 112, p. 214 (1926).

† Holborn and Olto, ' Z. f. Physik,' vol. 23, p. 77 (1924).

‡ Verschoyle, ' Roy. Soc. Proc.,' A, vol. 111, p. 552 (1926). (I am very much indebted to Mr. R. H. Fowler for drawing my attention in conversation to this paper and for pointing out its possible theoretical importance. The present paper is the outcome of that conversation.—J. E. L.-J.)

§ ' Roy. Soc. Proc.,' A, vol. 103, p. 524 (1923).

accounted for by the theory. The narrow range of temperature of the experiments is hardly sufficient for the further application of the results to determine intermolecular fields. Tentative calculations have, however, been made of the field of force between a hydrogen molecule and a nitrogen molecule, but it is hoped that the experimental data will soon be extended in a way which will permit of more accurate calculations.

The work of Masson and Dolley is devoted to a detailed examination of the deviations from Dalton's partial pressure law when the gases ethylene, argon and oxygen are mixed in various proportions. Unfortunately, the investigation is carried out at one temperature only, and this restriction prevents any application of the results in this paper.

An interesting property of the pressure of gaseous mixtures emerges from the investigation. It is found that even at the same temperature the pressure of a gaseous mixture may be greater than that of an equal concentration of either constituent alone. An explanation of this fact is offered, and it is shown for what proportions of the two gases the pressure of a mixture is a maximum.

2. *Theoretical Equation of State of a Gaseous Mixture.*

The equation of state of gas is usually expressed for theoretical purposes in the form

$$pv = k\mathrm{NT}\left(1 + \frac{\mathrm{B}}{v}\right), \tag{2.01}$$

so that the second term represents the deviation from the perfect gas law when the density is not too great. The coefficient B, which, following Kamerlingh Onnes, we shall refer to as the *second Virial Coefficient*, is a function of temperature and of the forces exerted by the molecules of the gas on each other. It is easy to evaluate it in terms of the law of force between molecules.* If this be of a spherically symmetrical type, so that it can be represented by a function of the distance only, $f(r)$, the formula for B proves to be

$$\mathrm{B} = 2\pi\mathrm{N}\int_0^\infty r^2\,(1-e^{2j\pi(r)})\,dr, \tag{2.02}$$

where $2j = 1/k\mathrm{T}$, k being the usual gas constant $(1{\cdot}372\,.\,10^{-16})$ and T the temperature, N the total number of molecules in the gas, and $\pi(r)$ the potential of the field given by

$$\pi(r) = -\int_0^\infty f(r)\,dr. \tag{2.03}$$

* 'Proc. Camb. Phil. Soc.,' vol. 22, p. 105 (1924).

We may also write equation (2.01) in the form

$$p = k\nu \mathrm{T} \{ 1 + \mathrm{B}'\nu \}, \tag{2.04}$$

where

$$\mathrm{B}' = 2\pi \int_0^\infty r^2 \{ 1 - e^{2j\pi(r)} \} \, dr, \tag{2.05}$$

and ν is the molecular concentration; we note that B′ is a function of the temperature which depends on the law of force.

If another gas be introduced into the same vessel, it is not difficult to show, by the same methods which have been used to give (2.05), that the resultant pressure is given by

$$p = k\mathrm{T} \{ (\nu_1 + \nu_2) + \nu_1^2 \mathrm{B}_{11}' + 2\,\nu_1\nu_2 \mathrm{B}_{12}' + \nu_2^2 \mathrm{B}_{22}' \}, \tag{2.06}$$

where

$$\mathrm{B}_{rs}' = 2\pi \int_0^\infty r^2 \{ 1 - e^{2j\pi_{rs}(r)} \} \, dr. \tag{2.07}$$

In a mixture of gases, the pressure is clearly given by

$$p = k\mathrm{T} \Sigma (\nu_r) + k\mathrm{T} \Sigma \, \nu_r\nu_s \mathrm{B}_{rs}' \tag{2.08}$$

where, in the second summation, r and s both range from 1 to n, and

$$\mathrm{B}_{rs}' = B_{sr}'.$$

When the B′s all vanish, the equation (2.08) reduces to the statement of Dalton's Partial Pressure Law. In actual gases, however, the coefficients B_{rs}' do not vanish, and so we have here a more general partial pressure law than that given by Dalton. If $p_1, p_2, \ldots p_n$ are the partial pressures of n gases in a given volume v, the total pressure p is given by

$$p = \Sigma (p_r) + k\mathrm{T} \Sigma' \, \nu_r\nu_s \mathrm{B}_{rs}', \tag{2.09}$$

where the accented summation sign indicates that squared terms are to be omitted. In the case of two gases only, we get simply

$$p = p_1 + p_2 + 2k\mathrm{T}\, \nu_1\nu_2 \mathrm{B}_{12}'. \tag{2.10}$$

The deviation from the Dalton Law thus gives us one simple method of determining B_{12}' and so of investigating the nature of the forces between the molecules of gas 1 with those of gas 2, though we shall not pursue this particular method here.

§ 3. *Observations of the Isotherms of Binary Mixtures and Comparison with Theory.*

In order to compare theory and experiment, it is necessary to represent the experimental results by an equation of the same type as (2.01). This was

first done by Kamerlingh Onnes,* who used a more general expansion, involving higher powers of $1/v$. These additional terms are appreciable, however, only at high pressures (of the order of 100 to 200 atmospheres), when it is unlikely that the theory applies. At lower pressures the observations can in most cases be sufficiently represented by an equation

$$pv = A + B/v. \tag{3.01}$$

In practice there is some divergence in the use of units in this virial equation. Leiden chooses as unit pressure the international atmosphere and takes the volume to be unity under unit pressure at 0° C. With these units, we may write equation (3.01) in the form

$$pv = A_v + \frac{B_v}{v}. \tag{3.02}$$

The symbol v then really represents a ratio of the volume to the volume under standard conditions. At Berlin, on the other hand, the unit of pressure is taken to be equivalent to a column of mercury 1 metre long (under standard conditions), and the unit of volume is chosen to be the normal volume under this pressure at 0° C. The equation may then be differentiated from that above by writing

$$pv = \mathfrak{A}_v + \frac{\mathfrak{B}_v}{v}. \tag{3.03}$$

Still another method of presenting the results has been used by Smith and Taylor† at Boston, but the equation used involves the assumption that the pressure varies linearly with the temperature at constant volume, which is not legitimate except over a small range. In any case, the equation is unsuitable for comparison with theory.

Again, it has proved convenient both at Berlin and at Leiden to express the value of pv, not in powers of $1/v$, but in powers of p, so that then we have

$$pv = A_p + B_p p \tag{3.04}$$

and

$$pv = \mathfrak{A}_p + \mathfrak{B}_p p. \tag{3.05}$$

The various coefficients are easily related to each other. Thus we find

$$A_p = A_v = \frac{\mathfrak{A}_p}{(\mathfrak{A}_p)_0 + (\mathfrak{B}_p)_0 l + (\mathfrak{C}_p)_0 l^2}, \tag{3.06}$$

$$B_p = \frac{B_v}{A_v} = \frac{l\mathfrak{B}_v}{(\mathfrak{A}_p)_0 + (\mathfrak{B}_p)_0 l + (\mathfrak{C}_p)_0 l^2}, \tag{3.07}$$

* Kamerlingh Onnes, 'Comm. Phys. Lab. Leiden,' No. 71, or 'Proc. Sect. of Sciences, Amsterdam,' vol. 4, p. 125 (1902).

† 'Journ. Amer. Chem. Soc.,' vol. 45, p. 2107 (1923).

where l is the pressure of one atmosphere in the Berlin units, and $(\mathfrak{A}_v)_0$ refers to the isothermal 0° C.

For comparison with the theoretical work, the expression (3.02) or (3.04) is probably the best, as it is referred to the normal conditions for which Avogadro's number for the molecular concentration is applicable. *These equations, therefore, we shall regard as the standard experimental equations to which all others are convertible.*

Theoretically, the equation (3.02) is more intelligible in terms of molecular concentration, so that if ν_0 is the concentration under the standard conditions and ν that under any other conditions, we have*

$$p = A_v \frac{\nu}{\nu_0} + B_v \frac{\nu^2}{\nu_0^2}, \tag{3.08}$$

the influence of temperature being represented by the variation of A_v and B_v with temperature.

For a simple gas, as has been mentioned in the above paragraph (equation (2.04)), theory gives an equation of state of the type

$$p = k\nu T\,(1 + B'\nu), \tag{3.09}$$

and this equation, being expressed in powers of ν, as equation (3.08) is, is directly comparable with the experimental results. Thus we deduce

$$k\nu_0 T = A_v, \qquad \nu_0 B' = \frac{B_v}{A_v}. \tag{3.10}$$

Since B' is a function of temperature which depends on the intermolecular forces, the second of these equations provides a criterion as to the suitability of any proposed model of the molecular field. It has been used in former papers to determine the molecular fields of the inert gases, on the assumption that these can be represented by the superposition of two fields, one repulsive and the other attractive, and both expressible as inverse power laws.†

From the theoretical equation of state of a binary mixture given in the preceding paragraph (2.06), we infer, as in (3.10), that

$$A_v = k\nu_0 T,$$

where ν_0 now refers to the concentration of the *mixture* under standard con-

* It is to be remembered that the unit of volume v is the volume of the gas under unit pressure at 0° C., so that $v = \nu_0/\nu$.

† 'Roy. Soc. Proc.,' A, vol. 106, p. 463 (1924); vol. 107, p. 157 (1925); vol. 112, p. 214 (1926).

ditions. As the $B'\nu^2$ of the single gas is replaced by the quadratic expression $B_{11}'\nu_1^2 + 2B_{12}'\nu_1\nu_2 + B_{22}'\nu_2^2$, we find also

$$\nu_0 (B_{11}'x_1^2 + 2B_{12}' x_1x_2 + B_{22}'x_2^2) = \frac{B_v}{A_v} = B_p, \tag{3.11}$$

or

$$B_{11}x_1^2 + 2B_{12}x_1x_2 + B_{22}x_2^2 = \frac{B_v}{A_v} = B_p, \tag{3.111}$$

where

$$B_{rs} = \nu_0 B_{rs}' \text{ and } x_1 = \nu_1/\nu,\ x_2 = \nu_2/\nu,\ x_1 + x_2 = 1. \tag{3.12}$$

We should therefore expect to find that the numerical value of B_v/A_v (or B_p) depends not only on temperature but also on the relative concentrations of the two gases. This is, in fact, the case. The experiments of Verschoyle* on mixtures of hydrogen and nitrogen at 0° and 20° C. have determined the values of B_v/A_v as a function of the relative concentrations, and these are plotted for the two temperatures 0° and 20° C. in fig. 1.†

Verschoyle does not give any theoretical explanation of the variation of B_v/A_v thus found, but contents himself with pointing out that the variation is *not linear*, as other observers seem to have expected. Thus the work of Amagat on oxygen and nitrogen is quoted, where the variation of B was found to be approximately linear. Now the condition that the left-hand side of (3.111) shall be linear in either x_1 or x_2 is that

$$B_{11} + B_{22} - 2B_{12} = 0.$$

This condition implies that the forces between any two molecules, like or unlike, are the same. This may be approximately true in the cases of oxygen and nitrogen, but will not be so generally.

In fig. 1 are included two curves, quadratic in ν_1/ν_2, which are drawn so that the mean square of the distances of the "observed" points from them are minima. It is clear that a quadratic relation of this kind satisfactorily accounts for the facts.

Moreover, the close agreement which is obtained justifies us in deducing from the observations the numerical values of each of the coefficients B_{11}, B_{12} and B_{22} in equation (3.111).

The values of the B's obtained in this way for the two temperatures are given in Table I, under Method 1. For comparison, the values of the B's obtained by drawing a quadratic curve through the two end points and the

* *Loc. cit.*

† Actually Verschoyle expresses his results in the form $pv = A_p + B_p p$ and determines A_p and B_p; but as seen from equation (3.07), B_p is the same as B_v/A_v.

middle point are also given in columns 3 and 4. These values do not differ greatly from those obtained by the more accurate method, and the corresponding

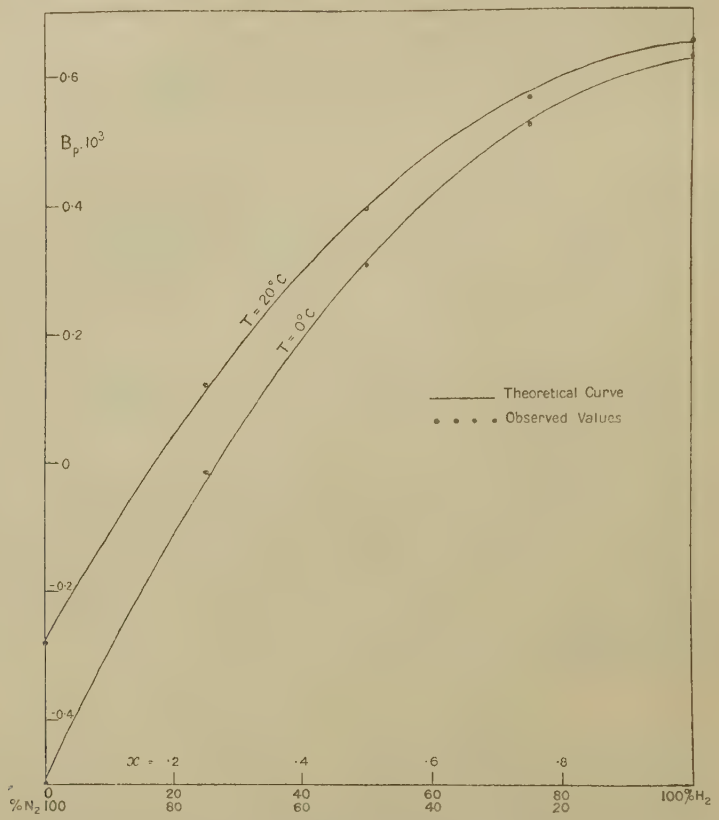


FIG. 1.—The theoretical curve for the second virial coefficient of a mixture and the observed values of Verschoyle for a mixture of H_2 and N_2.

graph of B is so near to that shown in fig. 1 that it is impossible to include it in the same figure.

Table I.—The " Partial " Virial Coefficients of a Hydrogen-Nitrogen Mixture.

	Method 1.		Method 2.	
	0° C.	20° C.	0° C.	20° C.
$B_{11} \cdot 10^3$	0·6224	0·6461	0·6263	0·6505
$B_{12} \cdot 10^3$	$0·5585_5$	0·6097	0·5483	$0·6104_5$
$B_{22} \cdot 10^3$	−0·4903	−0·2755	−0·4961	−0·2798

This assignment of numerical values to each of the B's individually is important theoretically, for the function B_{12} depends only on the forces between unlike molecules, just as B_{11} and B_{22} depend on the forces between like molecules. We are thus led naturally to a consideration of the forces between the various pairs of molecules. This we consider in detail in the subsequent paragraphs.

4. *The Determination of Intermolecular Forces.*

In the discussion of the preceding paragraph no special assumption was necessary about the nature of the intermolecular fields, but in order to make further progress it is necessary to particularise the law of force. We have given in equation (2.07) an expression for B'_{12}, etc., in terms of its potential field on the supposition that the molecular fields are spherically symmetrical. We now suppose that the force between two molecules can be represented by two inverse power laws

$$f_{12}(r) = \frac{(\lambda_n)_{12}}{r^{n_{12}}} - \frac{(\lambda_m)_{12}}{r^{m_{12}}}, \tag{4.01}$$

the first term to represent the repulsion, the second the attraction. A complete specification of the law of force then requires a knowledge of the force constants $(\lambda_n)_{12}$, $(\lambda_m)_{12}$, n_{12} and m_{12}.

This law of force has been used before for a single gas* and the integral in equation (2.07) has been evaluated. The corresponding expressions for B_{12}, etc., are

$$B_{12} = \tfrac{2}{3}\pi N_0 \left\{\frac{(\lambda_n)_{12}}{n_{12}-1} \cdot \frac{m_{12}-1}{(\lambda_m)_{12}}\right\}^{3/n_{12}-m_{12}} F(y_{12}) \tag{4.02}$$

where

$$y_{12} = \frac{(\lambda_m)_{12}}{kT(m_{12}-1)} \left\{\frac{(n_{12}-1)\,kT}{(\lambda_n)_{12}}\right\}^{\frac{m_{12}-1}{n_{12}-1}} \tag{4.03}$$

and

$$F(y) = y^{3/n-m}\left\{\Gamma\left(\frac{n-4}{n-1}\right) - \sum_{q=1}^{\infty} f(q)\, y^q\right\}. \tag{4.04}$$

In this function

$$f(q) = \frac{3\Gamma\left(\frac{q\overline{m-1}-3}{n-1}\right)}{q!\,(n-1)}. \tag{4.05}$$

The above expressions determine B_{12} as a theoretical function of the temperature when the constants of the molecular field are known, and, for a true representation of the field, this function should be identical with the value of B_{12} deduced from experiment (as described in the preceding paragraph).

* 'Roy. Soc. Proc.,' A, vol. 106, p. 463 (1924).

If the experiments cover a sufficiently wide range of temperature (as in Holborn and Otto's experiments on Neon,* −183° C. to 400° C.), the theoretical and experimental variation of B_{12} with temperature can be compared by graphical methods, as described in former papers.† The forces between unlike molecules can then be determined independently of the forces between the like molecules.

The experimental results on mixtures available at present are not sufficiently extensive to follow this method, and so recourse must be had to special methods which will make the best of the material to hand.

§ 5. *Comparison with Experimental Results.*

(1) *The Isotherms of a Helium-Neon Mixture by Holborn and Otto.*‡—Holborn and Otto have determined the isotherms of a mixture of helium and neon of fixed proportions (72·39 per cent. neon and 27·61 per cent. helium) over the temperature range 0° C. to 400° C. At the same time they determined the isotherms of pure helium alone, and later those of pure neon alone, in each case over the range 0° C. to 400° C. In each case they represented the results by the law

$$pv = \mathfrak{A}_v + \mathfrak{B}_v p.$$

We have three series of results of the type

$$x_1^2 B_{11} + 2x_1x_2B_{12} + x_2^2B_{22} = B_v/A_v$$

corresponding to $x_1 = 1$, $x_2 = 0$; $x_1 = 0{\cdot}7239$, $x_2 = 0{\cdot}2761$; $x_1 = 0$, $x_2 = 1$, in each case the right-hand side being an observed function of temperature.§ The first and the last series determine B_{11} and B_{22} as functions of temperature, and so B_{12} is determined as a function of the temperature by the equation—

$$B_{12} = \frac{1}{2a_1a_2}\left\{\left(\frac{B_v}{A_v}\right)_{12} - a_1^2\left(\frac{B_v}{A_v}\right)_{11} - a_2^2\left(\frac{B_v}{A_v}\right)_{22}\right\}, \qquad (5.01)$$

where $a_1 = 0{\cdot}7239$, $a_2 = 0{\cdot}2761$. This can then be dealt with in the same way as the observed B of a single gas. The values for the various temperatures are given in Table II.

The values of log B_{12} can now be plotted against log T and the graph compared with the theoretical graph of log $F_{12}(y_{12})$ against log y_{12}, given by

* 'Z. f. Physik,' vol. 33, p. 1 (1925).

† Lennard-Jones and Cook, 'Roy. Soc. Proc.,' A, vol. 112, p. 214 (1926), § 2.

‡ *Loc. cit.*, 'Z. f. Physik,' vol. 23, p. 77 (1924).

§ The values of B_v/A_v are obtained from the values of A_p and B_p, given by Holborn and Otto, in the way described in §3.

Table II.—The values of B_{12} for a Helium-Neon Mixture (Holborn and Otto).

T.	log B_{11} (Neon)	log B_{22} (Helium).	log $\left(\frac{B_v}{A_v}\right)$ Ne–He.	log (B_{12}).
0°	$\bar{4}\cdot6770$	$\bar{4}\cdot7228$	$\bar{4}\cdot7239$	$\bar{4}\cdot7787$
100°	$\bar{4}\cdot7232$	$\bar{4}\cdot7054$	$\bar{4}\cdot7512$	$\bar{4}\cdot7930$
200°	$\bar{4}\cdot7648$	$\bar{4}\cdot6931$	$\bar{4}\cdot7614$	$\bar{4}\cdot7689$
300°	$\bar{4}\cdot7880$	$\bar{4}\cdot6702$	$\bar{4}\cdot7672$	$\bar{4}\cdot7558$
400°	$\bar{4}\cdot7869$	$\bar{4}\cdot6548$	$\bar{4}\cdot7513$	$\bar{4}\cdot7179$

equations (4.04) and (4.03) respectively. The method of determining the most suitable molecular model and the numerical values of the force constants has been described in previous papers* and need not be repeated here.

Table III.—The Intermolecular Fields of Helium and Neon.

n.	m.	—	He – He.	He – Ne.		Ne – Ne.	$\frac{\sigma_{11}+\sigma_{22}}{2}$.
				(1).	(2).		
9	5	$\lambda_n \cdot 10^{74}$	1·14	2·54	2·93	6·32	—
		$\lambda_m \cdot 10^{45}$	4·35	8·68	10·00	22·3	—
		$\sigma \cdot 10^{8}$	4·03	4·45	4·53	4·99	4·51
11	5	$\lambda_n \cdot 10^{89}$	0·625	1·35	1·76	4·38	—
		$\lambda_m \cdot 10^{45}$	2·91	5·20	6·83	17·2	—
		$\sigma \cdot 10^{8}$	3·53	3·82	3·92	4·29	3·91
$14\frac{1}{3}$	5	$\lambda_n \cdot 10^{115}$	2·15	4·253	5·49	22·2	—
		$\lambda_m \cdot 10^{45}$	1·90	2·39	3·28	13·3	—
		$\sigma \cdot 10^{8}$	3·10	3·27	3·33	3·70	3·40

The five experimental points available are shown in fig. 2, together with the theoretical curve, corresponding to $n = 11$, $m = 5$. This figure is a sufficient indication of the inadequate range of the experimental results. As the value of B_{12} for the isothermal of 400° C. does not appear to be consistent with the rest, two attempts have been made to use the results, (1) by considering the 5 points as shown, (2) by considering only 4 points. The results of these two methods are given in Table III, and for comparison the forces between helium and helium, neon and neon are included from a former paper.†

As in other papers a convenient measure of the repulsive forces is a quantity

* 'Roy. Soc. Proc.,' A, vol. 106, p. 463 (1924); vol. 107, p. 157, § 2 (1925); vol. 112, p. 214, § 2 (1926).

† Lennard-Jones and Cook, *loc. cit.*

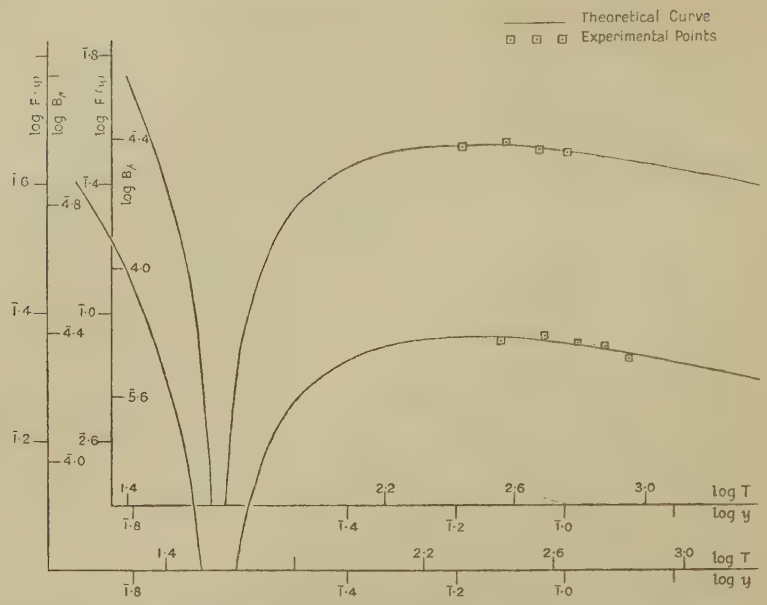


FIG. 2.—The Experimental Values of the Second Virial Coefficient of a Helium-Neon Mixture and the Theoretical Curve.

$\sigma_{rs}^{(n)}$ which has the dimensions of a length and corresponds to the diameter of a molecule when it is assumed to be rigid. It is defined by the relation

$$\sigma_{rs}^{(n)} = \left\{\frac{2\,(\lambda_n)_{rs}}{3\,(n-1)\,k}\right\}^{\frac{1}{n-1}}. \qquad (5.02)$$

The calculated values for helium and neon are given in Table III.

(2) *The Isotherms of a Hydrogen-Nitrogen Mixture by Verschoyle.**—In paragraph 3 we deduced from Verschoyle's work on hydrogen and nitrogen the values of B_{11}, B_{12}, and B_{22} for 0° C. and 20° C. The limitation of the experiments to these two temperatures prevents us from applying the graphical methods just described. In this particular case, therefore, we consider a somewhat different method, applicable when the B's are known at two temperatures only. Let these temperatures be T_1 and T_2, and let the corresponding values of y, which are related to the temperature by equation (4.03) be denoted by y_1 and y_2. Then, on using equation (4.02), we find

$$\frac{B_{rs}(T_1)}{B_{rs}(T_2)} = \frac{F(y_1)}{F(y_2)} \qquad (5.03)$$

and

$$\frac{y_1}{y_2} = \left(\frac{T_2}{T_1}\right)^{\frac{n-m}{n-1}}. \qquad (5.04)$$

* *Loc. cit.*

The last equation determines y_2 in terms of y_1 and so, on substituting in equation (5.03), we are left with a polynomial in y_1 of the form

$$(1-\gamma_{rs})\,\Gamma\left(\frac{n-4}{n-1}\right) - \sum_{q=1}^{\infty}\left\{1-\gamma_{rs}\left(\frac{T_1}{T_2}\right)^{\frac{n-m}{n-1}q}\right\} f(q)\, y_1{}^q = 0 \qquad (5.05)$$

where

$$\gamma_{rs} = \frac{B_{rs}(T_1)}{B_{rs}(T_2)}\left(\frac{T_1}{T_2}\right)^{3/n-1} \qquad (5.06)$$

and $f(q)$ has been defined in equation (4.05). If, therefore, we assign definite values to n and m, the exponents of the repulsive and attractive fields, equation (5.05) determines y_1. From this value, we can deduce the appropriate force constants from the equations

$$\left\{\frac{(\lambda_n)_{rs}}{(\lambda_m)_{rs}}\right\}^{3/n-m} = \frac{3B_{rs}(T_1)}{2\pi\nu_0\left(\frac{m-1}{n-1}\right)^{3/n-m} F(y_1)} \qquad (5.07)$$

and

$$(\lambda_m)_{rs}^{\frac{n-m}{n-1}}\left\{\frac{(\lambda_m)_{rs}}{(\lambda_n)_{rs}}\right\}^{\frac{m-1}{n-1}} = \frac{m-1}{(n-1)^{\frac{m-1}{n-1}}}(kT_1)^{\frac{n-m}{n-1}} y_1 \qquad (5.08)$$

The first of these determines the ratios of the force constants, and when this result is used in the second equation, the individual value of $(\lambda_m)_{rs}$ is determined. The results of the calculations are given in Table IV.

This method has the disadvantage that it does not indicate which of an assumed set of values of n and m are to be preferred. For this information the temperature variation of the B's over a wide range is necessary.

Table IV.—The Intermolecular Forces of Hydrogen and Nitrogen.

n.	m.	—	$H_2 \longleftrightarrow H_2$.	$H_2 \longleftrightarrow N_2$.	$N_2 \longleftrightarrow N_2$.
9	5	$\lambda_n \cdot 10^{73}$	1·10 (0·919)*	2·74	19·17 (15·82)
		$\lambda_m \cdot 10^{44}$	2·90 (2·54)	5·40	19·94 (18·20)
		$\sigma \cdot 10^{8}$	5·35 (5·23)	5·99	7·64 (7·46)
11	5	$\lambda_n \cdot 10^{89}$	8·85 (7·38)	25·24	298·9 (224·5)
		$\lambda_m \cdot 10^{44}$	2·20 (1·98)	4·14	16·73 (15·10)
		$\sigma \cdot 10^{8}$	4·61 (4·52)	5·12	6·55 (6·36)
$14\frac{1}{3}$	5	$\lambda_n \cdot 10^{114}$	5·74 (4·89)	22·6	559·1 (361·0)
		$\lambda_m \cdot 10^{44}$	1·67 (1·56)	3·27	13·9 (12·3)
		$\sigma \cdot 10^{8}$	3·97 (3·92)	4·40	5·60 (5·42)

* The numbers in brackets were obtained in a former paper (*loc. cit.*) from the more extensive experimental work of Holborn and Otto (*loc. cit.*).

The results for hydrogen and nitrogen have each been treated by this method as well, although their molecular fields have been investigated previously by

the method described above.* The results of the latter calculations are included in Table IV (in brackets) for purposes of comparison. The agreement between the two sets of results is as good as could be expected in view of the small temperature difference between the isothermals from which the calculations of this section have been made. For this reason too much reliance must not be placed on the present calculations. The former calculations are likely to be much more reliable. The calculations have been made here as a practical illustration of the methods suggested. The theory now awaits more extensive experimental results.

§ 6. *Discussion of the Results.*

One of the chief points of interest is the relation of $(\lambda_n)_{12}$, the repulsive force constant between a molecule 1 and a molecule 2, to $(\lambda_n)_{11}$ and $(\lambda_n)_{22}$. In a former paper an empirical rule has been adopted for lack of any evidence. It was there suggested that, as the relation is

$$\sigma_{12} = \tfrac{1}{2}(\sigma_{11} + \sigma_{22})$$

when the molecules are actually rigid spheres, the same relation would probably be true for the generalised diameter.†

$$\sigma^{(n)} = \left(\frac{2\lambda_n}{3\,(n-1)\,k}\right)^{1/n-1}$$

leading to the relation

$$\lambda_{12}^{1/n-1} = \tfrac{1}{2}(\lambda_{11}^{1/n-1} + \lambda_{22}^{1/n-1}).$$

The results given in Table III for the forces between helium and neon atoms show how far this expression is satisfied. In the last column are included the values of $\frac{1}{2}(\sigma_{11} + \sigma_{22})$, which are to be compared with the values of the σ given in the column headed He – Ne. It is seen that in the case of the values obtained by the second method there is very close agreement, especially when $n = 9$ and $n = 11$. When more experimental information is available about the isothermals of a helium-neon mixture, so that the force can be obtained with more confidence, the above assumed relationship can be put to a stricter test. The results for the hydrogen-nitrogen mixture cannot be used for a test of this kind for reasons which have already been stated.

* Lennard-Jones and Cook, *loc. cit.*

† Lennard-Jones and Taylor, ' Roy. Soc. Proc.,' A, vol. 109, p. 483 (1925).

§ 7. *Properties of a Gaseous Mixture.*

An examination of the experimental results of Holborn and Otto on helium and neon shows that when the isothermals are represented in each case by the formula

$$pv = \mathfrak{A}_p + \mathfrak{B}_p p$$

the numerical value of $\mathfrak{B}_p$ for the mixture is in some cases greater than those of the separate gases. For this reason the actual results of Holborn and Otto are here collected from their papers* and reproduced in Table V.

Table V.—Values of the Virial Coefficient $\mathfrak{B}_p$ from Holborn and Otto.

$\mathfrak{B}_p \cdot 10^3$	Helium.	Helium–Neon.	Neon.
0°	0·6954	0·6970	0·6257
100°	0·6680	0·7424	0·6960
200°	0·6493	0·7600	0·7660
300°	0·6160	0·7702	0·8080
400°	0·5945	0·7425	0·8060

Since $\mathfrak{B}_p$ is small, the above equation can be written

$$\frac{pv}{\mathfrak{A}_p} = 1 + \frac{\mathfrak{B}_p}{v}$$

where v is, in each case, the ratio of the volume to the volume under standard conditions; it is therefore inversely proportional to the molecular concentration. Since $(\mathfrak{B}_p)_{12}$ is greater than $(\mathfrak{B}_p)_{11}$ and $(\mathfrak{B}_p)_{22}$ at the temperatures 0° C. and 100° C., it follows that at these temperatures the pressure of the gaseous mixture is greater than that of either constituent alone, even when the molecular concentration is in every case the same. This is a natural consequence of the fact that the virial coefficient has a maximum. In the case of helium this maximum occurs at a temperature of — 140° C. about, while in the case of neon the maximum occurs at a temperature in the neighbourhood of 260° C. The maximum value of $\mathfrak{B}_p$ for the helium-neon mixture lies between these temperatures, with the result that the value of $\mathfrak{B}_p$ for the mixture is greater for a certain range of temperature than the corresponding values for the constitutents.

The theoretical formula for B given in this paper and elsewhere† also exhibits a maximum, as fig. 2 shows. It is the only theoretical formula with this property which has yet been given.

* 'Z. f. Physik,' vol. 23, p. 86 (1924); vol. 33, p. 5 (1925).

† *Cf.* the curves given in 'Roy. Soc. Proc.,' A, vol. 107, p. 157, figs. 2 and 3 (1925).

From equation (3.111) it is easily seen that the virial coefficient will have its maximum value (at a given temperature) when the relative concentration of the gases is given by

$$\frac{x_1}{x_2} = \frac{B_{12} - B_{22}}{B_{12} - B_{11}}.$$

When the fields of force of the gases have once been determined, the numerical values of B_{11}, B_{12} and B_{22} can be calculated for any temperature from the formula given in paragraph (4), and the ratio of x_1 x_2 for a maximum value of B_p can then be predicted. This concentration would give the greatest pressure for the same total number of molecules in the gas.

§ 8. *Summary.*

(1) A theoretical formula is given for the equation of state of a gaseous mixture.

(2) This formula shows the deviation from Dalton's Partial Pressure Law.

(3) It is shown that the second virial coefficient of a binary mixture is a quadratic function of the relative concentrations. This law satisfactorily accounts for some recent experimental results.

(4) A method of determining the forces between the unlike molecules of a mixture is deduced.

(5) A special feature of the pressure of a binary gaseous mixture is pointed out and a theoretical explanation is given.

The Influence of Boundary Films on Corrosive Action.

By L. H. Callendar, Ph.D., A.R.C.S., B.Sc.

(Communicated by H. C. H. Carpenter, F.R.S.—Received March 22, 1927.)

Introduction.

Now that the electrochemical character of corrosion has been firmly established,* it becomes necessary to examine in detail those electrolytic factors which most influence corrosive action in normal waters.† It is the purpose of this paper to consider the most important of these factors, namely, the presence of boundary films on the metal formed both before and after contact with the electrolyte. The rapid formation of films of moisture, grease, etc., on a fresh surface exposed to the air has been pointed out by the late Lord Rayleigh.‡ Our knowledge of the properties of these films on metals has been greatly extended by the work of Hardy,§ but their influence on the location of corrosion on a metal surface does not appear to have been investigated. The boundary films formed between the electrodes and the electrolyte as a result of the passage of a current have recently been investigated by Newbery,|| and it would appear that these films must also have an important influence on corrosive action.

It is generally admitted that the rate of corrosion in normal waters is partly dependent on the oxygen supply. But before the metal comes in contact with the electrolyte it may have undergone oxidation, and it seems that the importance of surface condition, and the profound effect it may have on the reactivity of the metal has not hitherto been fully realised, for in many previous researches on corrosion, often insufficient attention has been given to the method of preparing the specimens. The results obtained in this paper appear to indicate that the primary cathode and anode areas on a pure metal are determined largely by the distribution of oxide and other films on the metal surface. When the metal is immersed in an electrolyte the first current flows between these primary areas, but subsequently the location of the cathode and anode areas may be altered, as Evans¶ has shown, by variations in the oxygen concentra-

* W. D. Bancroft, ' J. Phys. Chem.,' vol. 28, p. 868 (1924).
† Waters of Ph 4·5—11 in Sorensen's logarithmic notation.
‡ Lord Rayleigh, ' Collected Papers,' vol. 3, p. 523.
§ W. B. and J. K. Hardy, ' Phil. Mag.,' vol. 38, p. 32 (1919).
|| ' Roy. Soc. Proc.,' A, vol. 111, p. 182 (1926).
¶ ' J. Inst. Metals,' vol. 30, p. 239 (1923).

tion, by the spreading of the products of corrosion over the metal surface, and by other factors.

Lambert* has suggested that with a pure metal the original potential differences over the surface are due to lack of homogeneity and strains within the metal, but the P.D. caused by these factors is usually extremely small and may be negligible in cases where exposure to air has resulted in the formation of certain kinds of films. Similarly the presence of noble impurities within the metal will seldom be a serious cause of corrosion in normal waters, for although the P.D. may be large the cathode area is likely to be extremely small, and will not be nearly so favourable for oxygen depolarisation as the large area provided by the normal oxide film on the metal surface.

During the course of the electrolytic action films of a gaseous, liquid, or solid nature are formed between the electrodes and the electrolyte, and these films generally tend to slow down the rate of corrosion. Investigations have been carried out by Newbery and others as to the nature and electrical resistance of these films, and H. D. Holler† has used the term " boundary resistance " to signify the resistance of such films. In the present paper it is proposed to use the term " boundary resistance " to signify the resistance in ohms of the boundary films between the electrodes and the electrolyte, obtained by subtracting the resistance of the electrolyte (measured by Kohlrausch's method) from the total internal resistance of the experimental corrosion cell; this boundary resistance is an indicator of the rate of corrosion.

The experimental corrosion cells used in this research consisted of an aluminium anode and a platinum or oxidised aluminium cathode with various dilute salt solutions as electrolytes. Aluminium forms the most convenient and sensitive metal for the electrodes, because, in addition to its tendency to form highly resistant scales, the solution potential of the oxidised metal differs widely from that of the metal itself, and therefore any action between the aluminium and the electrolyte is readily indicated by changes in the resistance and total e.m.f. of the corrosion cell.

The electrical method used for measuring the changes in resistance and e.m.f. was described in a previous paper by the author,‡ and a further detailed description with various improvements in the adaptability and accuracy of the method is given in a later section of this paper. The method is primarily

* ' Trans. Chem. Soc.,' vol. 101, p. 2056 (1912).

† ' U.S. Bur. Stand.,' Scientific Paper No. 504 (1925).

‡ " Passification and Scale Resistance in Relation to the Corrosion of Aluminium Alloys," 'J. Inst. Metals,' vol. 34, p. 57 (1925).

designed to indicate changes in resistance and e.m.f., and the results must be interpreted accordingly, but its extreme simplicity and adaptability seem to make it very suitable for the investigation of many other aspects of corrosion besides those dealt with in this paper.

The Metal Used for Experiments.

Three different samples of aluminium metal were used and these will be referred to in experiments by the letters A, P, and S.

Sample A was a very pure American metal* kindly supplied to the author by Dr. H. J. Vernon of the Non-Ferrous Metals Research Association. Sample P was a normal pure commercial metal. Sample S consisted of a test-piece containing large single crystals made by the process introduced by Carpenter and Elam† and kindly supplied to the author by Miss Elam ; part of one large single crystal was cut out and used as an electrode.

The analysis of these three materials is given in Table I. Samples A and P were also examined spectroscopically,‡ and the spectrum photograph showed traces of Mg. Ni in sample P ; P also appeared to contain a minute trace of carbon.

Table I.—Chemical Analysis of Metal used in Experiments. (Per cent.)

Sample.	Iron.	Silicon.	Copper.	Titanium.	Sodium.	Aluminium.
A*	0·017	0·014	0·014	—	—	Remainder
P	0·150	0·200	—	—	—	—
S	0·073	0·125	—	0·018	0·016	—

* Chemical analysis by D. P. Bayley.

Other impurities indicated by the spectrum analysis but not determined chemically should not exceed 0·010 per cent. in the case of any one element. Both samples A and P were in the form of rolled sheets one-quarter of an inch thick.

The electrodes made from the aluminium sheets P and A were six in number and all of the same size, 3 by 2 by $\frac{1}{4}$ inch. They were stamped with the letter of the alloy and the serial number thus : A_1, A_2, P_1, P_2, P_3, P_4. As only one surface of the electrode was required to take an active part in the experiments,

* For method of manufacture see J. D. Edwards, ' Amer. Electrochem. Soc.,' April, 1925.

† ' Roy. Soc. Proc.,' A, vol. 100, p. 329 (1921).

‡ ' Spectrum Analysis,' by Prof. H. Dingle.

the back and sides were coated with an insulating layer consisting of aluminium paint covered with spar varnish, and finally coated before each experiment with white paraffin wax; this was found to give an adherent insulating coating having no effect on the conductivity of the electrolyte. The electrodes were fitted with easily removable screw terminals, and fixed at a distance of 1 inch apart by means of detachable ebonite distance pieces as shown in fig. 1.

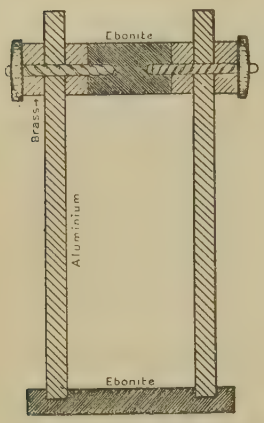


FIG. 1.—Pair of Aluminium Electrodes.

For the preparation of the salt solutions used in the electrolytic cells two samples of conductivity water were used. The first sample was kindly supplied to the author by Mr. J. M. Stuart, from a special still constructed by the Corrosion Research Department at the Royal School of Mines; as delivered from the still it had the exceptionally low conductivity of $0{\cdot}08\frac{1}{4} \times 10^{-6}$ reciprocal ohms, and was almost entirely free from dissolved salts, carbon dioxide, and ammonia, etc. The second sample of water prepared by distillation over alkaline permanganate, the first portion containing ammonia being rejected, was used in some comparative experiments. Before use both samples were shaken with the laboratory air, which was free from the usual acid and alkaline fumes, and it was found impossible to detect any difference between them sufficient to affect the results.

The Salts used and the Conductivity of their Solutions.

The salts used were of the standard of the British Drug Houses analytical reagents. The solutions for use as electrolytes in the corrosion cells were made up in various strengths from one-thousandth to one-tenth normal. For the purpose of this paper it was necessary to know the resistance of the electrolyte in the corrosion cells, and therefore measurements were made of the conductivities of these electrolytes as actually used in the experiments. For preliminary determinations the usual Kohlrausch method was used, a cell being fitted with platinised platinum electrodes, and connected to a source of alternating current and a telephone detector. After some preliminary results had been obtained a special cell was substituted for the laboratory one, and this new cell had electrodes of the same size (3×2), and at the same distance apart (1 inch), as those used in the standard corrosion cell. After the resistance of different electrolytes had been measured in this cell, the small

platinum cathode used in a number of the special corrosion experiments was substituted for one of the electrodes, and thus a further series of electrolyte resistances were obtained. These measurements were carried out in a thermostat at 20° C., the same temperature being used in the corrosion experiments. The figures so obtained for the resistance of the electrolytes are given in the experiments that follow and are used in the calculation of the boundary resistances.

Apparatus and Method of Experiment.

The arrangement of the apparatus used in these experiments for the measurement of changes in the resistance and total e.m.f. of the electrolytic cell is shown in fig. 2.

The electrolytic cells, c_1, c_2, c_3, were placed in the thermostat which was usually maintained at a temperature of 20° C. Three experiments were

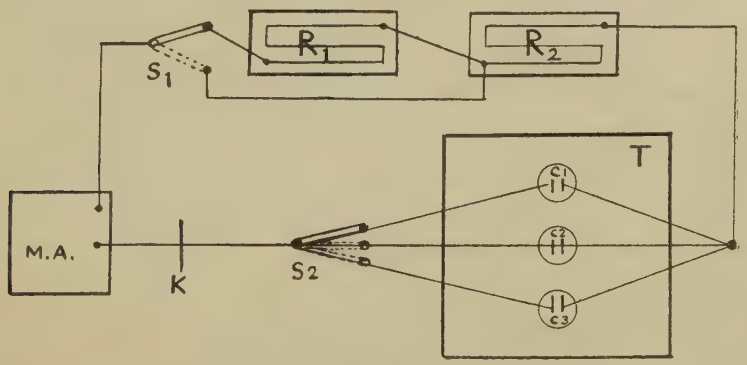


Fig. 2.

generally carried on at the same time. The cells were connected to the switch S_2, and the microammeter (M.A.) through the key (K), and on the other side to the two resistance boxes and the switch S_1 which serves to put one or both of the boxes into the circuit and connect them with the microammeter. The electrolytic cell consisted of a 250 c.c. beaker filled with the electrolyte to such a depth that the liquid just covered 2 inches of the electrodes (this required 160 c.c. of electrolyte), the latter being placed in the cell just before taking the first reading. The microammeter used for measuring the current in the circuit had an internal reistance of 500 ohms. In experiments where the polarisation is changing it is important to take readings quickly without waiting for each deflection of the instrument to become steady ; in such cases the limit of the first swing of the needle was observed. The corresponding values of the steady current were obtained from an auxiliary table, based on the results of a preliminary calibration of the instrument, in which the final steady deflection was

observed in addition to the first swing. With the two resistance boxes the external resistance can be varied from 43,000 to zero ohms.

Calculation of the Resistance and Total e.m.f. of the Cell.

In the actual experiments measurements were made of the current given by the cell when its terminals were connected through a high resistance R_1 and then through a lower resistance R_2, and it was assumed for purposes of calculation that, during the time required to take the reading (3 to 4 seconds), the total e.m.f. of the cell did not alter. This assumption was tested by measuring the rate of change of the total e.m.f. with different external resistances and different electrolytes using a cell with aluminium anode and platinum cathode. and it was found that the maximum change in the calculated total e.m.f. in the time required to take one reading (using the same electrolytes as in the actual experiments) was less than 1 per cent.; a total error of this order is evidently quite immaterial in these comparative experiments. The results of experiments were therefore calculated from the following formula :—

If R_1 and R_2 are the total external resistances,
and i_1 and i_2 the measured currents,
then let E be the total e.m.f.,
and R_0 the internal resistance of the cell.

We have $(R_1 + R_0)\, i_1 = E$

$(R_2 + R_0)\, i_2 = E$,

whence $R_1 i_1 - R_2 i_2 = R_0 (i_2 - i_1)$,

therefore R_0 = internal resistance of cell $= \dfrac{R_1 i_1 - R_2 i_2}{i_2 - i_1}$.

From the resistance R_0 so obtained the boundary resistance may be calculated by subtracting the resistance of the electrolyte measured by Kohlrausch's method, and the total e.m.f. E of the cell is readily deduced. It will be seen that these figures, when considered in conjunction with the conditions of the experiment and the time of immersion of the electrodes, serve to indicate what factors retard or accelerate corrosive action.

Order of Experiments.

The experiments in Section A show how the surface potential of aluminium alters with different treatment, and suggest that the rise in potential corresponds to the covering of the surface with oxide. The distribution of this oxide is dependent upon the surface conditions of the metal, and the experiments in Section B show that this determines the reactivity of the metal.

The experiments in Section E show that whether the cathode film is formed on the metal in air, or in the electrolyte, or by heating, it will, in a dilute chloride solution, readily maintain a current with the polished metal. The final experiments F are introduced to show the influence of these oxide films on the boundary resistance and rate of corrosion.

Experimental Section A.

In the following experiments a number of aluminium electrodes were treated in various ways, and their potential was then measured against a platinum cathode in a sodium chloride electrolyte.

Experiments 2600–04.—Effect of previous surface treatment on the potential of pure aluminium in a dilute sodium chloride solution.

Electrodes.—Aluminium anode and small platinum cathode heated before each reading.

Electrolyte.—N/1000 sodium chloride solution.

Resistance of electrolyte—1,549 ohms.

External resistance, $R_1 = 43{,}500$ ohms. $R_2 = 3{,}500$ ohms.

Experiment No.	Electrode.	Surface treatment.	Total cell e.m.f.
			mvs.*
2600	P_1	Polished on 000 Hubert paper under paraffin and then rinsed with pure benzene to remove grease (see W. B. and J. K. Hardy, *ibid.*)	1,207
	P_2		1,228
			1,204
			1,228
			1,235
	A_2	Ditto ditto	1,158
2602	P_1	Polished as above and then stood in conductivity water for 4 minutes	1,031
	A_2	Polished and stood in conductivity water for 18 hours	727
	P_2	Polished and stood in the laboratory air for 18 hours	910
2604	P_1	Polished on a clean dry selvyt	1,080
		As above, but selvyt wetted with conductivity water	941
	P_1	Polished on wet selvyt with alumina	875
	A_2		877
2603	P_4	Polished electrode (30 at 170°)	886
	P_2	Heated in air for (50 at 170°)	851
	P_2	Time minutes (60 at 185°)	842
	P_2	Temp. ° C. (120 at 170°)	854
	P_1	Time hours (65 at 300°)	936

* Note mvs. = mille-volts.

The above figures tend to show that there is a difference of potential of some 300 mvs. between clean aluminium freshly polished in air and the same

metal that has stood in the air, been heated in the air, or immersed in pure water, or polished with alumina.*

When aluminium is heated in air Pilling and Bedworth† have shown that an oxide film forms on its surface the thickness of which depends on the temperature and time of heating, it therefore seems justifiable to assume from the above results that about 300 mvs. roughly represents the P.D. between an aluminium surface freshly formed in air and the oxide of the metal. It is possible that the solution potential of the dry oxide may alter slightly with the temperature of formation, and also the dry oxide probably differs somewhat from the moist oxide or hydroxide formed in water (the low figure 727 mvs., obtained where the metal had been immersed for a long time in water, may be due to the absorption of hydrogen by the film); it also appears that an oxide film can be obtained by polishing with alumina. The next experiment shows the rate at which this rise of potential of aluminium takes place in air at ordinary temperatures.

Experiment 2610.—Rate of change of surface potential in air.‡

Electrodes.—S_1, polished on 0,00,000 Hubert paper under paraffin and washed with benzene. Exposed to air for the times indicated in column T in the results below; re-polished and cleaned for each exposure. For the readings the aluminium electrode is connected to the small platinum cathode (previously depolarised by heating) in an N/1000 sodium chloride electrolyte.

Resistance of electrolyte = 2,340 ohms.

R_1 = 43,500 ohms. R_2 = 3,500 ohms.

Time of exposure to air. T.	Measured cell resistance.	Boundary resistance.	Total e.m.f.	Fall in total e.m.f.
	ohms.	ohms.	mvs.	mvs.
Zero	3,900	1,560	1,230	—
35 minutes	3,900	1,560	1,090	140
55 ,,	4,105	1,765	1,082	148
100 ,,	4,210	1,870	1,050	180
120 ,,	4,110	1,770	1,017	213
15 hours	4,595	2,255	923	307

The rate of fall of the total e.m.f. is plotted in fig. 3.

* McAuley and Bowden ('J. C. S.,' vol. 127, p. 2605 (1925)) say that there are two normal states in which iron and zinc tend to exist, and the P.D. between these two states is for zinc = 75, iron = 200 mvs.

† 'J. Inst. Met.,' vol. 29, p. 529 (1923).

‡ G. C. Schmidt ('Z. Phys. Chem.,' vol. 106, p. 105 (1925)) has made comparisons of the solution potentials of a number of metals both before and after grinding with emery. He finds that grinding produces a lowering of the potential and he has measured

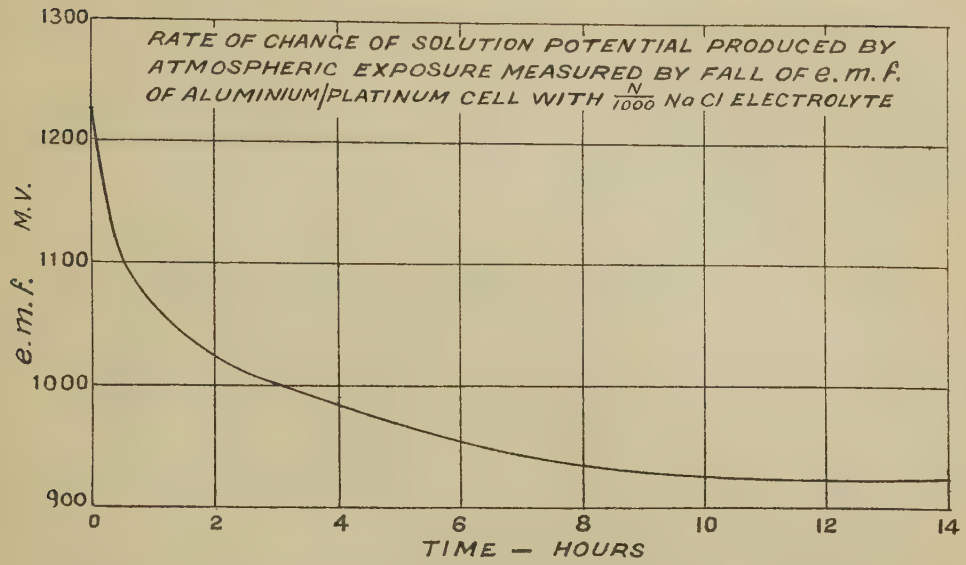


FIG. 3.

Section B.—Effect of Surface Condition on Reactivity.

The object of the following experiments is to show how greatly the initial reactivity varies with the condition of the metal surface, and to suggest that surface condition, by determining oxygen distribution, locates the primary cathode and anode areas on the metal surface. The surfaces compared in an aluminium-platinum cell with dilute oxidising electrolyte are: (1) turned surface, (2) clean emery polished surface, (3) greasy emery polished surface, (4) alumina polished surface, (5) pitted and rough ground surface.

Experiment 2516.—Turned surface.

Electrolyte.—N/1000 potassium dichromate solution.

Electrodes.—Anode A_1 turned on a lathe with a fine tool. Cathode, small platinum ($1{\cdot}4 \times 0{\cdot}5$ inch) heated to a white heat and allowed to cool before connecting up for each reading.

Resistance of electrolyte = 1,490 ohms.

R_1 = 42,500 ohms, R_2 = 2,500 ohms.

Time.	Measured resistance.	Boundary resistance.	Total e.m.f.
	ohms.	ohms.	mvs.
Zero	not measured		
34 minutes	3,500	2,010	1,140
75 ,,	3,790	2,300	1,120
151 ,,	3,920	2,430	1,072
233 ,,	4,300	2,810	1,068
18 hours	5,080	3,590	773

the rate at which the various metals recover their normal potential (this "normal" potential would probably be found to correspond closely to that of the normal oxide of the particular metal); Schmidt attributes the differences of solution potential so obtained to surface tension effects.

Experiment 2518.—Clean emery polished surface.

Electrolyte.—As in experiment 2516.

Electrodes.—Anode P_1 polished on 000 Hubert polishing paper under paraffin and then rinsed with pure benzene. Cathode small platinum as in experiment 2516.

External resistance as before.

Time.	Measured resistance.	Boundary resistance.	Total e.m.f.
	ohms.	ohms.	mvs.
Zero	not measured		
6 minutes	3,570	2,080	1,106
15 ,,	3,880	2,390	1,050
34 ,,	4,400	2,910	988
70 ,,	4,610	3,120	894
126 ,,	5,350	3,860	868

Experiment 2511.—Greasy emery polished surface.

Electrolyte.—Similar to experiment 2518.

Electrodes.—Anode P_2 polished as in experiment 2518 under paraffin but not washed with benzene, *i.e.*, surface left greasy. Other conditions exactly as in experiment 2518.

Time.	Measured resistance.	Boundary resistance.	Total e.m.f.
	ohms.	ohms.	mvs.
Zero	not measured		
4 minutes	4,520	3,030	1,263
14 ,,	5,000	3,510	1,169
21 ,,	5,070	3,580	1,156
38 ,,	5,320	3,830	1,150
157 ,,	6,330	4,840	1,032

Experiment 2516A.—Alumina polished surface.

Electrolyte.—N/1000 potassium dichromate solution.

Electrodes.—Anode A_2 first polished and cleaned as in experiment 2518, and then finished with fine alumina powder and water on polishing wheel. Cathode as before.

Time.	Measured resistance.	Boundary resistance.	Total e.m.f.
	ohms.	ohms.	mvs.
Zero	not measured		
7 minutes	4,250	2,760	856
17 ,,	4,720	3,230	820
36 ,,	5,170	3,680	766
77 ,,	6,030	4,540	701
153 ,,	6,710	5,220	634
235 ,,	7,300	5,810	591
18 hours	7,600	6,110	406

Experiment 2512.—Pitted surface.

Electrolyte.—As before.

Electrodes.—Anode P_4 heavily pitted by immersion in pure neutral hydrogen peroxide solution for one week in a vessel open to the air, washed with conductivity water, and then rubbed over with coarse emery paper. Platinum cathode as usual.

Time.	Measured resistance.	Boundary resistance.	Total e.m.f.
	ohms.	ohms.	mvs.
Zero	not measured		
2 minutes	4,840	3,350	1,156
5 ,,	5,410	3,920	1,130
11 ,,	6,270	4,780	1,112
20 ,,	6,870	5,380	1,073
136 ,,	8,720	7,230	952

Comments on Series B Experiments.

The resistance and e.m.f. changes obtained in these experiments are plotted in figs. 4 and 5 against the time of immersion. The increase of resistance in

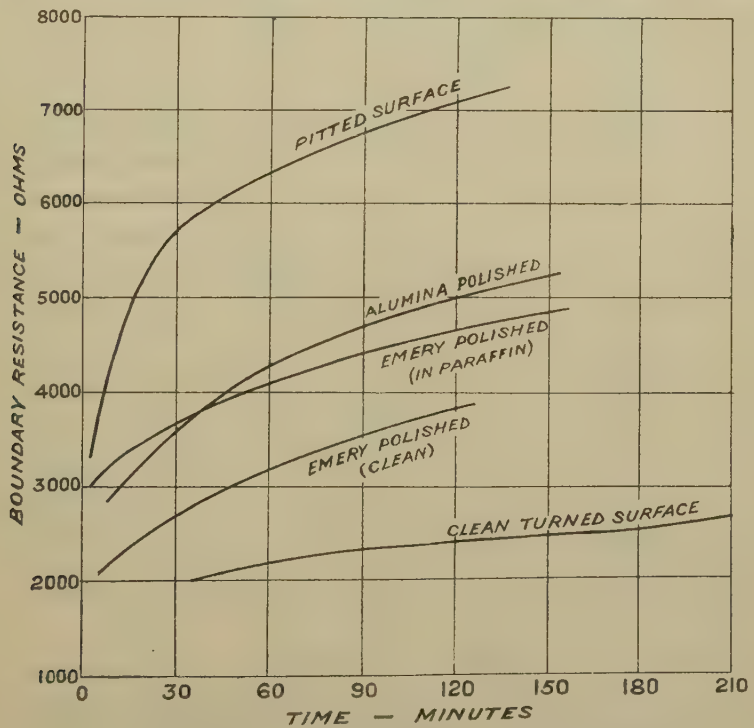


Fig. 4.—Effect of Surface Condition on Boundary Resistance.

each case indicates the formation of boundary films due to electrolytic action between anodes and cathodes on the surface of the aluminium, that is to say,

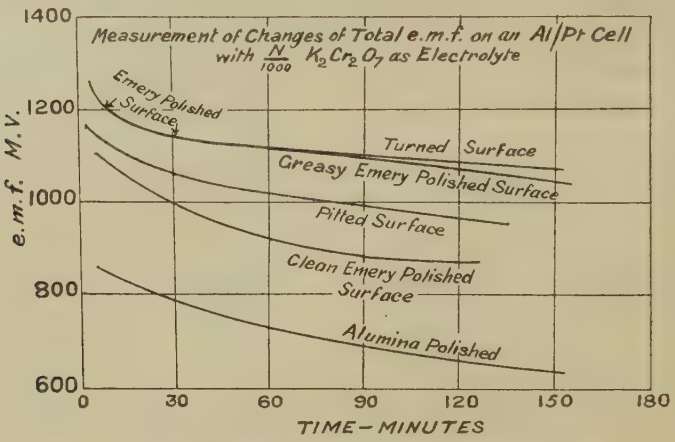


FIG. 5.—Effect of Surface Condition on Solution Potential.

that under the conditions of these experiments the *rate* of increase in the boundary resistance appears to indicate tendency to corrosion. The clean-turned and the emery-polished surfaces are the least reactive, the emery-polished surface being more reactive than the turned surface partly because it contains imbedded in the metal numerous emery grains, which can easily be seen under the microscope, and because it no doubt contains a certain amount of cathodic oxide below the surface which would be completely removed by the deeper cut of the turning operation ; the greasy emery polished surface and the alumina polished surface are very similar in reactivity—the one is reactive owing to grease and emery particles isolating anode areas and the other owing to uneven distribution of alumina over the surface ; the old pitted surface is, as might be expected, the most reactive.

The curves in fig. 5 show the potential changes, and while the initial values differ somewhat, the most noteworthy feature is the very low e.m.f. value for the cell with an alumina polished anode. This low value is evidently due to a film of oxide which is probably formed by the alumina polishing powder being pressed into the soft aluminium surface ; a special experiment on the effect of alumina-polishing the surface of a single crystal is given in Section F, and this tends to confirm the above suggestion.

Experimental Series, E.—Illustrating conditions favouring Local Corrosion.

Having shown in the Series A experiments that the surface of aluminium is rendered more cathodic by oxidation in the air or by immersion in a dilute

electrolyte, it is now necessary to illustrate by experiment the kind of conditions which favour the maintenance of a current between scale-covered metal and the clean metal. Since the presence of chlorides in a water is known to be favourable to the local corrosion of metals, we have therefore used a dilute sodium chloride solution as electrolyte in the model corrosion cell experiments which follow ; further, the area of the cathode used is much larger than that of the anode, a condition favouring cathodic depolarisation by oxygen ; finally, oxygen from the air has ready access to the open beaker in which are placed the electrodes partly immersed in the electrolyte.

One electrode is made cathodic by (1) exposure to the air, (2) heating in the air, (3) immersion in the electrolyte, and the cathode is connected to a clean polished aluminium anode. While the cell is giving a continuous external current through a high resistance, readings are taken by the usual method, and the changes in the boundary resistance at the electrodes and the total e.m.f. of the cell are obtained.

Experiment 2651.—Cathode film formed in air.

Electrodes.—P_2 polished on 0,00,000 Hubert emery paper under paraffin, rinsed with benzene, and then exposed to the laboratory air for 16 hours before being connected to the freshly polished P_3 and immersed in the electrolyte, an N/100 sodium chloride solution.

Size of electrodes.—$P_2 = 2$ by 2 inch immersed, $P_3 = 2 \times \frac{1}{2}$ in.

Resistance of electrolyte = 157 ohms.

$R_1 = 20{,}500$ ohms, $R_2 = 500$ ohms.

Time.	Boundary resistance.	Total e.m.f.
	ohms.	mvs.
Zero	1,286	250
2 minutes	3,003	142
8 ,,	4,153	129
19 ,,	5,043	113
40 ,,	5,593	113
76 ,,	7,483	129

Experiment 2650.—Cathode film formed by heating.

Electrodes.—P_2 polished and cleaned as in experiment 2651, and then heated in an air oven for 2 hours at 190° C., cooled and connected to freshly polished P_3 in a similar sodium chloride electrolyte; the same external resistances were used.

Time.	Boundary resistance.	Total e.m.f.
	ohms.	mvs.
Zero	2,528	221
3 minutes	5,003	154
17 ,,	5,203	148
38 ,,	6,388	131

Experiment 2657.—Cathode film formed by heating.

Electrodes.—A_2 polished and cleaned as in experiment 2651, and then heated in an electric oven for 16 hours at 300° C., cooled and connected to freshly polished P_3.

Electrolyte.—N/250 sodium chloride solution.

Resistance of electrolyte = 372 ohms.

R_1 = 23,500 ohms, R_2 = 3,500 ohms.

Time.	Measured resistance.	Boundary resistance.	Total e.m.f.
	ohms.	ohms.	mvs.
Zero	6,000	5,628	288
2 minutes	12,200	11,828	172
6 ,,	14,200	13,828	91
18 ,,	19,200	18,828	56

Polished and cleaned both A_2 and P_3 and reconnected in same electrolyte. Current reverses, A_2 now the anode, and P_3 the cathode, reading:—

	ohms.	ohms.	mvs.
Zero	900	528	57

Experiment 2658.—Cathode film formed by immersion.

Electrodes.—A_2 polished and cleaned as in experiment 2651, immersed in the chloride electrolyte for 100 minutes, and then connected to the freshly polished P_3.

Electrolyte.—N/250 sodium chloride solution.

Resistance of electrolyte and the external resistances as in experiment 2657.

Time.	Measured resistance.	Boundary resistance.	Total e.m.f.
	ohms.	ohms.	mvs.
Zero	4,270	3,898	311
3 minutes	4,140	3,768	188
18 ,,	9,990	9,618	147
68 ,,	8,050	7,678	76
147 ,,	8,530	8,158	77

Results of Experiments, Series E.

The previous experiments show that when the metal is rendered cathodic by exposure to air, by heating, or by immersion in an electrolyte it will in a dilute chloride solution maintain a current with the polished metal, and it would equally appear that a current should be set up if part of a metal surface were rendered cathodic by any of the above processes, the rest of the surface remaining anodic.

Experiments F, on Resistant Films.

As experiment 2657 shows a large increase in the boundary resistance when a thick film is present on the cathode, it is important to investigate further the effect of such films on the rate of corrosion. In the following experiments a comparison is made between the film-covered and the polished metal and the influence of such films on the rate of corrosion is indicated.

Experiment 2701.—Oxide film on metal.

Electrolyte.—N/500 sodium chloride solution.

Electrodes.—Small platinum cathode heated before each reading. Anode A_2 polished under paraffin with 0,00,000 Hubert emery papers and rinsed in benzene; heated in an air oven for 6 hours at 160° C., allowed to cool, and then connected in the electrolyte.

Resistance of electrolyte = 800 ohms.

R_1 = 43,500 ohms, R_2 = 3,500 ohms.

Time.	Measured resistance.	Boundary resistance.	Total e.m.f.
	ohms.	ohms.	mvs.
Zero	4,490	3,690	828
4 minutes	4,860	4,060	890
13 ,,	4,660	3,860	903
29 ,,	5,000	4,200	905
79 ,,	5,320	4,520	890
180 ,,	5,200	4,400	864

Experiment 2712.—Oxide film on metal.

Electrolyte.—N/1000 sodium chloride solution.

Electrodes.—Small platinum cathode heated as usual. Anode P_1 polished as in experiment 2701, and then heated in an electric air oven for 65 hours at 300° C.

Resistance of electrolyte = 1,549 ohms.

R_1 = 43,500 ohms, R_2 = 3,500 ohms.

Time.	Measured resistance.	Boundary resistance.	Total e.m.f.
	ohms.	ohms.	mvs.
Zero	9,930	8,381	936
3 minutes	13,630	12,081	1,000
10 ,,	14,400	12,851	1,014
35 ,,	14,200	12,651	975
114 ,,	13,450	11,901	996
18 hours	14,000	12,451	954

Experiment 2711.—Oxide film on metal.

Electrolyte.—N/250 sodium chloride solution.

Electrodes.—Platinum cathode as before. Anode P_2 polished as in experiment 2701 and heated for 65 hours at 300° C.

Resistance of electrolyte = 402 ohms.

External resistance as in experiment 2712.

Time.	Measured resistance.	Boundary resistance.	Total e.m.f.
	ohms.	ohms.	mvs.
Zero	4,850	4,448	924
5 minutes	7,900	7,498	974
11 ,,	8,310	7,908	972
23 ,,	8,160	7,758	976
107 ,,	8,160	7,758	976
273 ,,	7,500	7,098	951
18 hours	7,510	7,108	905

Experiment 2720.—Clean polished metal.

Electrolyte.—N/1000 sodium chloride solution.

Electrodes.—Platinum cathode as usual. Anode A_2 polished as in experiment 2701, connected with the cathode and at once immersed in the electrolyte.

Resistance of electrolyte = 1,549 ohms.

External resistance as before.

Time.	Measured resistance.	Boundary resistance.	Total e.m.f.
	ohms.	ohms.	mvs.
Zero	3,200	1,651	1,142
4 minutes	4,820	3,271	1,092
15 ,,	4,420	2,871	875
32 ,,	4,760	3,211	854

Experiment 2721.—Clean polished metal.

Electrolyte.—N/250 sodium chloride solution.

Electrodes.—Platinum cathode as usual. Anode A_2 polished as in experiment 2701, connected with the cathode and immersed in the solution.

Resistance of electrolyte = 402 ohms.

External resistance as before.

Time.	Measured resistance.	Boundary resistance.	Total e.m.f.
	ohms.	ohms.	mvs.
Zero	not measured		
2 minutes	2,175	1,773	1,003
6 ,,	2,710	2,308	954
27 ,,	2,900	2,498	901
18 hours	3,640	3,238	834

The results of the preceding five experiments are plotted on the curves figs. 6 and 7.

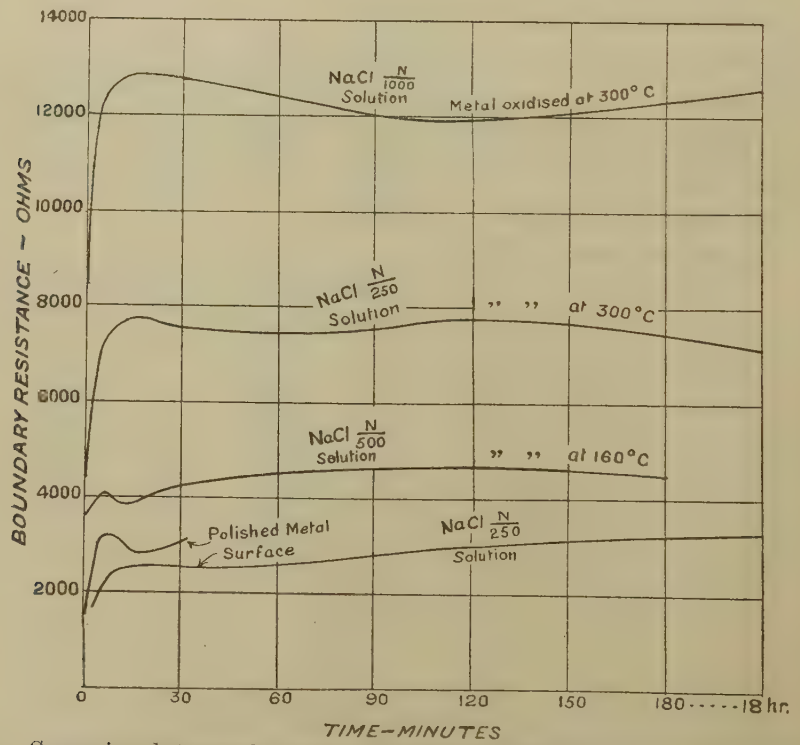


Fig. 6.—Comparison between the Action of Sodium Chloride Solution on an Oxidised and on a Polished Aluminium Surface. Boundary Resistance in an Aluminium/Platinum Cell with Sodium Chloride as Electrolyte.

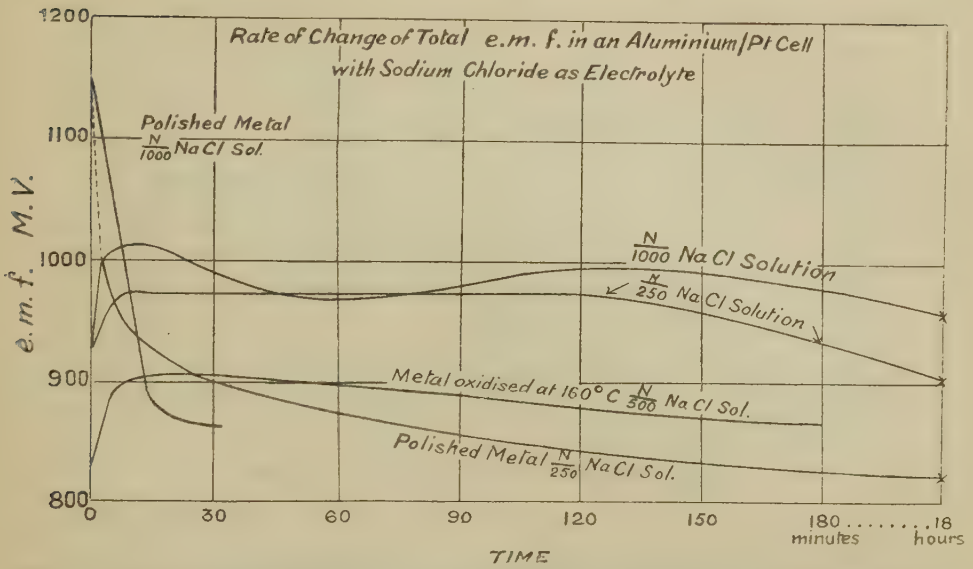


FIG. 7.—Comparison between the Action of Sodium Chloride on an Oxidised and on a Polished Aluminium Surface.

Results of Experimental Series F on Boundary Oxide Film.

The curves in fig. 6 show that where the metal is already covered with an oxide film a high boundary resistance is obtained, and that this resistance increases with the thickness of the film (a thicker film being presumedly formed at 300° C. than at 160° C. in the same time), and with increasing dilution of the electrolyte, The curves in fig. 7 for e.m.f. changes show an initial rise of e.m.f. when the oxide film is present, and a rapid fall where no film was originally present.

This increase in boundary resistance with film thickness is evidently a very important factor in corrosion, and in Table III below some interesting comparative figures are given.

While the initial boundary resistance may be mainly due to gas layers on the metal surface, the increasing resistance shown in the last three experiments of Table III is mainly due to the presence of solid films of increasing thickness.

Table III.—Influence of an Oxide Film on the Boundary Resistance. Electrolyte, N/1000 NaCl solution. Aluminium Anode, Platinum Cathode.

Experiment No.	Condition of metal surface.	Initial boundary resistance.
		ohms.
2720	Clean emery polished surface	1,651
2602	Film formed by 18 hours' exposure to air at 20° C.	2,501
2700	Oxide film formed by heating at 160° C. for 8 hours	4,511
2712	Oxide film formed by heating at 300° C. for 65 hours	8,381
2721	Visible film—Anodic oxidation*	26,300

* The electrolyte used here was N/500 NaCl solution. For method of preparation of these films see "The Anodic Oxidation of Aluminium and its Alloys as a Protection against Corrosion." H.M. Stationery Office, 1926, and British Patents Nos. 223994/5, by G. D. Bengough and J. M. Stuart.

Films Formed by Polishing with Alumina.

Since the rise in potential obtained by polishing aluminium with alumina indicates the formation of an oxide film, it is interesting to examine how a surface film so formed behaves in comparison with a polished surface; to emphasise the difference between such surfaces a sodium nitrate solution was used as electrolyte; for the experimental surface a single crystal of pure aluminium was used.

Experiment 2522.—Alumina-finished surface.

Electrolyte.—N/500 sodium nitrate solution.

Electrodes.—Anode S_1, single crystal of aluminium polished on Hubert paper under paraffin and finished by six hours' polishing on a rotating wheel with fine alumina powder and conductivity water; a fine bright reflecting surface was obtained. The anode was then rinsed in benzene to remove grease and connected to a recently heated platinum cathode and immersed in the electrolyte. The cathode was heated before each reading.

Resistance of electrolyte = 1,500 ohms.

R_1 = 43,500 ohms, R_2 = 3,500 ohms.

Time.	Measured resistance.	Boundary resistance.	Total e.m.f.
	ohms.	ohms.	mvs.
Zero	8,000	6,500	686
17 minutes	8,180	6,680	616
56 ,,	10,720	9,220	594
105 ,,	13,400	11,900	541
172 ,,	17,200	15,700	514
18 hours	21,600	20,100	251

Experiment 2526.—Clean emery polished surface.

Electrolyte.—N/500 sodium nitrate solution.

Electrodes.—Anode S_1 single crystal of aluminium polished on Hubert paper under paraffin, rinsed with benzene, and then connected to the platinum cathode; the latter being heated before each reading.

Electrolyte and external resistance as before.

Time.	Measured resistance.	Boundary resistance.	Total e.m.f.
	ohms.	ohms.	mvs.
Zero	2,790	1,290	1,051
15 minutes	4,390	2,690	840
31 ,,	5,240	3,740	830
61 ,,	6,070	4,570	775
18 hours	12,210	10,710	421

Comments on Alumina Polishing Experiments.

The alumina polished surface gives initially boundary resistance and e.m.f. figures approximately the same as the clean metal does after about 2 hours' immersion in the nitrate solution; this tends to indicate that both immersion in a nitrate electrolyte and alumina polishing result in the formation of an oxide film in close contact with the metal and similar to that formed by heating in air; in any case the film is strongly cathodic to the metal itself and would be likely to set up a current with any part of the metal from which it was removed.

Discussion of Results.

Boundary Films.—The first series of experiments indicate various ways in which the surface potential of aluminium may be altered, and this alteration in potential is attributed in the first place to the formation of an oxide film on the metal surface (though when the metal is immersed in an electrolyte there may be a further fall in the e.m.f. of the corrosion cell due to hydrogen polarisation). It is evident the commercial metal will be normally covered with an oxide film, but as shown in experiment 2610 this film forms relatively slowly in the air at ordinary temperatures, so that while it is forming its even distribution may be prevented by the settling on the metal surface of impurities from the air such as dust, grease, moisture, etc., and by irregularities in the surface itself; it therefore appears that the primary cathodic areas over the metal surface will be determined by this purely accidental distribution of oxide. The primary anodic areas, where local solution of the metal will start, will be those

parts which were covered in the air but which the electrolyte can reach, for example covered by some material soluble in, or easily washed off, or penetrated by the electrolyte such as salt, dust or sand, and those parts which were not oxidised in the air for any reason such as their being covered up by the metal resting on them, or because of their depth below the surface and their relative inaccessibility. The conclusion is therefore reached that the initial tendency to corrode and the distribution of corrosion are mainly determined by the primary boundary film on the metal.

Surface Condition and Reactivity.—Further confirmation of the above conclusion is obtained from the next experiments in which the effect of differences in the surface condition of the metal is emphasised by using an oxidising electrolyte in a special corrosion cell. First there is a clean turned surface which shows very little activity, for on immersion in the solution it becomes evenly oxidised and any slight differences in potential over the surface are equalised by the conveyance of bichromate ions to the anode areas, this very dilute solution being apparently sufficient for the purpose. The clean emery-polished surface is apparently more reactive than the turned surface, and this may be attributed to the presence of emery grains (dust from the air may be expected to have a similar accelerating action). The alumina-polished and greasy surfaces are more reactive owing in the one case to the uneven distribution of alumina, and in the other to the location of the anode areas being mainly determined by grease patches acting as oxygen screens. The old pitted surface, although roughly polished, shows by far the greatest reactivity for the relative inaccessibility of the bottom of the pits, and the old semi-porous films over the mouths prevent oxygen and bichromate ions from reaching the main anode areas, so that there is all the more oxygen available for hydrogen depolarisation and the maintenance of a large P.D. between the cathode areas and the anodic pits.

Scale—Metal Corrosion Cell.—The experiments in Section E show that the maintenance of a current between the scale-covered and polished metal in a dilute solution (chloride) is equally successful whether the cathode film is formed by exposure to air, by immersion in an electrolyte, or by heating in the air; in practice the metal will always be liable to exposure to one or other of these methods of oxidation, and when the metal comes in contact with an electrolyte, an uneven distribution of the oxide film may be the origin of local corrosion. In all these experiments there is a fall in the total e.m.f. of the cell, and this is mainly due to oxidation of the polished anode by oxygen in the water, for it is found that, if the anode is repolished, the total e.m.f. of the cell rises almost to

the original figure. It is seen, therefore, that the important factors for the continuance of local corrosion are, that oxygen should be prevented from reaching the anode areas and that the cathode areas should have a regular supply.

Influence of Oxide Films on the Rate of Corrosion.—It has been shown by Pilling and Bedworth* that the oxide film formed by heating aluminium in air continues to thicken even at a temperature of 600° C., until the metal has been heated for about 70 hours, but after this oxidation apparently ceases. Since the process of manufacture requires the heating and melting of the metal we may often expect to find the surface of the commercial metal covered with a thick oxide film ; it is, therefore, important to examine the effect of such films on the corrosion of the metal. Theoretically it appears that if the film is locally removed just before immersion in the electrolyte the exposed area will be anodic to the film and will corrode, but from the experiments in Section F it is seen that the film greatly increases the initial boundary resistance, therefore it appears that although some action might take place, the presence of the film must greatly reduce the rate of corrosion. These experiments show also that the boundary resistance increases both with the thickness of the film and with the increasing dilution of the electrolyte, which suggests that while part of the resistance may be that of the film, a part is also due to an adsorbed layer of spent electrolyte. The final experiments show that a similar resistant oxide film can be formed on aluminium by merely polishing with alumina and water.

In conclusion it seems probable that the results obtained in this paper must in many cases supply the explanation for the initiation of local corrosion of other metals besides aluminium, in fact it would appear that wherever a metal (*e.g.*, iron, zinc, etc.) tends to form oxide films or scales of higher potential than the metal itself, it may for this very reason be liable to local corrosion ; the boundary resistance effects will vary with different metals according to the particular properties of the films formed.

Summary.

The local corrosion of metals in normal waters is governed by boundary films of a solid, liquid, or gaseous nature present on the metal surface.

The surface of those metals liable to local corrosion is normally more or less covered with an oxide film, and where this film is of higher potential than the metal itself, its distribution determines the location of the primary cathode and anode areas before the metal comes in contact with the electrolyte. The

* N. B. Pilling and R. E. Bedworth, ' J. Inst. Metals,' vol. 29, p. 529 (1923).

distribution of this oxide film is determined by the presence of foreign substances on the metal surface and by irregularities in the surface itself.

When the metal comes in contact with the electrolyte the oxide film is the primary cathode and metal passes into solution at unoxidised parts of the surface, and experiments show that the continuance of this current between the cathode film and the metal is dependent on the prevention of diffusion of oxygen to the anodes, and they also indicate that the original location of the cathode and anode areas is likely to be altered by the distribution of oxygen within the solution.

The boundary resistance between the electrodes and the electrolyte is an indicator of the rate of corrosion; the normal cathodic oxide film formed in air is so thin that it has little effect on the rate of corrosion, but the thicker oxide films formed by heating give a high boundary resistance and must tend to retard corrosive action; oxidising electrolytes also retard corrosion by increasing boundary resistance.

With aluminium the boundary resistance increases both with increasing dilution of the electrolyte and with increasing thickness of any oxide film present on the metal surface.

The author wishes to acknowledge his indebtedness to Prof. H. C. H. Carpenter, F.R.S., for arranging for the work to be done in his laboratories and for the interest he has shown in its progress; and he also wishes to express his thanks to Mr. R. May and Mr. J. M. Stuart for reading through the paper and making valuable suggestions.

The Band Spectrum of Water Vapour.

By David Jack, M.A., B.Sc., Assistant and Carnegie Teaching Fellow, The University, St. Andrews.

(Communicated by O. W. Richardson, F.R.S.—Received March 25, 1927.)

[Plate 6.]

1. *Experimental Evidence on the Nature of the Emitter.*

Experiments on the spectrum of burning hydrogen or hydrogen compounds by Liveing and Dewar* led to the discovery of the now well-known band spectrum of water vapour in the ultra-violet. Its discovery was published simultaneously by Huggins.† The presence of water vapour, or of both its constituent elements in flames, in an electric discharge or in an arc, always gives rise to this spectrum, which consists of several well-marked bands. With oxygen, even when it has been subjected to the most powerful means of drying, it is not possible to eliminate completely the water-vapour bands, and consequently it was believed that these bands were due to oxygen.‡ Eder and Valenta§ showed, however, that the addition of a small quantity of hydrogen to the oxygen in a discharge tube greatly enhanced the intensity of the bands. Their appearance with dry oxygen is presumably due to the emission by the electrodes of small traces of hydrogen, which combine with the oxygen under the influence of the discharge. Grebe and Holtz,‖ in 1912, came to the conclusion that the emitter was some combination of oxygen and hydrogen, although not necessarily H_2O. Reiss¶ obtained the spectrum with cyanogen burning in very dry oxygen, and in 1914 Fortrat** attributed the spectrum to oxygen. More recent experiments have been carried out by Watson,†† who found that when water vapour was excited by a weak electrodeless discharge, the series lines α, β and γ of hydrogen made their appearance along with the water-vapour bands. This indicates that even with weak excitation the water

* ‘Roy. Soc. Proc.,’ vol. 30, p. 580 (1880); vol. 33, p. 274 (1882).
† ‘C. R.,’ vol. 90, p. 1455 (1880).
‡ W. Steubing, ‘Ann. d. Physik,’ vol. 33, p. 573 (1910).
§ ‘Beitrage zur Photochemie u. Spektralanalyse,’ p. 21.
‖ ‘Ann. d. Physik,’ vol. 39, p. 1243 (1912).
¶ ‘Z. f. Phys. Chemie,’ vol. 88, p. 555 (1914).
** ‘J. de Physique,’ vol. 5, p. 20 (1924).
†† ‘Astrophys. Journ.,’ vol. 3, p. 145 (1924).

molecule may be broken up, at least into OH and H. Under similar conditions he obtained the benzene bands without any evidence of the breaking up of the complex benzene molecule. When water vapour was subjected to a powerful disruptive discharge, the characteristic oxygen bands appeared, which showed that under these conditions complete dissociation of the water molecules was taking place. Watson concluded that the bands were probably due to the OH ion, and in support of this view he gave further evidence, which will be referred to later, based on an estimate of the moments of inertia of the emitter.

2. *Bands* λ 3064 *and* λ 3122.

Among the earlier measurements of the water-vapour spectrum may be mentioned those of Liveing and Dewar,* Meyerheim,† and Grebe and Holtz.‡ Some of the lines have been arranged into series by Meyerheim, and Deslandres and D. Azambuja,§ but to Heurlinger|| is due the credit of giving a complete arrangement into series of the lines measured by Grebe and Holtz. From these measurements Heurlinger picked out two bands with heads at λ 3064 and λ 3122. With the exception of a group of closely packed lines probably forming the head of the band 3122, for which region complete details are not given by Heurlinger, all but 10 of the 260 lines measured by Grebe and Holtz are included. Fortrat succeeded in resolving some of the lines given as single by Grebe and Holtz and gave the wave-numbers of 7 lines forming the head of the band 3122. Photograph No. 1 of Plate 6 shows the appearance of the bands 3064 and 3122. The band 3122 is much weaker than the band 3064, by which it is entirely overlaid, so that it is not obvious as a separate band in the photograph. In fact, Heurlinger considered these two as forming a single band, but the work of Dieke¶ and of the author shows that they must be treated as quite distinct.

3. *General Structure of the Water-Vapour Bands.*

Each band consists of six branches, for which the wave-numbers of the lines fit approximately into parabolic formulæ, which are, however, not in exact agreement with the elementary quantum theory of band spectra. Heurlinger denoted these branches P_1^p, P_2^p, Q_1^p, Q_2^p, R_1^p, R_2^p, where $p = 1, 2$. The index,

* 'Phil. Trans.,' A, vol. 179, p. 27 (1888).
† 'Z. f. wiss. Phot.,' vol. 2, p. 131 (1904).
‡ *Loc. cit.*
§ 'C. R.,' vol. 157, p. 814 (1913).
|| "Untersuchungen über die Struktur der Bandenspektra," Lund (1918).
¶ 'K. Akad. Wetensch. Amsterdam,' vol. 28, p. 174 (1925).

p, was used to distinguish the branches which belong to what are now considered as separate bands. The use of the index will therefore be dropped. The branches lie fairly closely together in pairs forming doublet branches, which will be denoted, P, Q, R. The structure of the bands may best be illustrated by plotting as abscissæ the wave-numbers, ν, of the lines and as ordinates the rotation quantum numbers, m. This is done in fig. 1 for the band 3064.

The components of the doublets are widely separated for the lower values

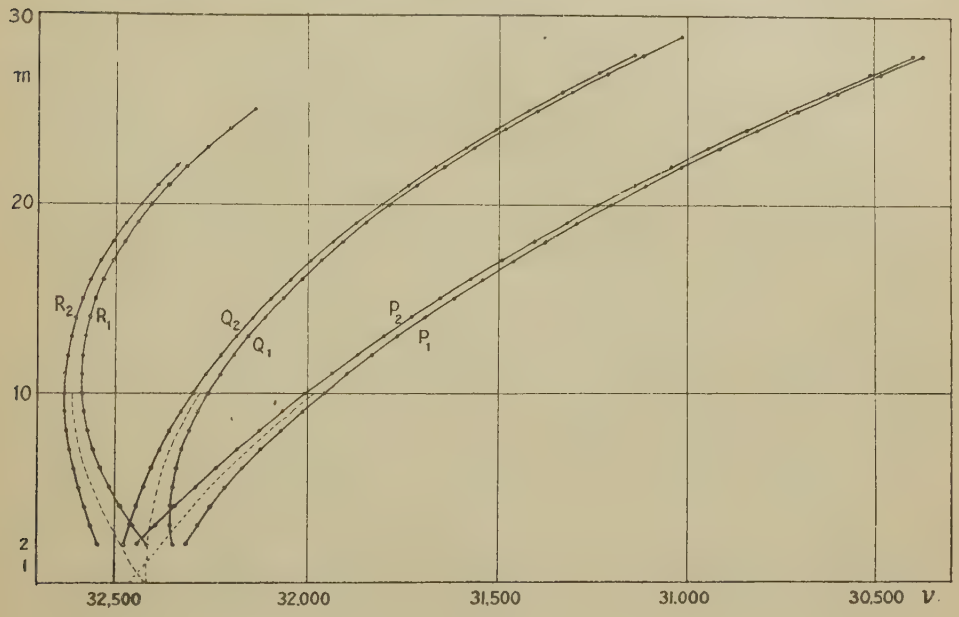


Fig. 1.

of m and become closer as m increases. Some of the doublets belonging to the P branch of the band 3064 are plainly visible in the photograph, Plate 6, No. 1.

Heurlinger gave several empirical relations between the doublet separations. He showed that, except for irregular variations,

$$P_2(m) - P_1(m) = R_2(m) - R_1(m) - 2\delta,$$

and

$$P_2(m) - P_1(m) = Q_2(m) - Q_1(m) - 2\varepsilon(m),$$

where δ is a constant and $\varepsilon(m)$ a function of m, and $P_1(m)$ represents the wave-number of the line in the P_1 branch having the quantum number m. The function $2\varepsilon(m)$ is very small for the lower values of m, but for the higher

values it amounts to about 20 per cent. of $P_2(m) - P_1(m)$. The separation of the components in any one branch is given by an expression of the form,

$$P_2(m) - P_1(m) = 2\phi(1/m),$$

where $\phi(1/m)$ is a power series in $1/m$. Heurlinger further pointed out that corresponding doublets in the two bands have almost equal separations, but that the deviation appears to be systematic.

4. *Bands* λ 2811 *and* λ 2875.

Watson* gave details of two bands, λ 2811 and λ 2875, which are similar in structure to the two already mentioned. He measured these bands from photographs taken in the second order of a 21-foot concave grating. They are much weaker than the bands measured by Grebe and Holtz, and exposures of six to eight hours were necessary to obtain satisfactory photographs, using a powerful discharge through water vapour. In order to eliminate as far as possible the continuous spectrum of hydrogen, which is troublesome when a discharge tube of the usual form is used, the apparatus was arranged so that a stream of water vapour could be kept flowing through the tube. Watson also made measurements of a series of single lines lying near the head of the band 3064. These lines fit fairly closely into a parabolic formula. The bands 2811 and 2875, and also the singlet series, are shown in photograph No. 2, Plate 6.

Comparison of the results given by Heurlinger and Watson reveals a striking similarity between the various bands. The separation in corresponding doublets is practically identical for the two bands 3064 and 2811, and the same is true of the bands 3122 and 2875. Watson estimated the moments of inertia of the emitter of the bands 3064 and 2811 and obtained the same final value in both cases, but appreciably different initial values. Dieke† showed that the differences, $Q_i(m) - P_i(m-1)$ and $R_i(m) - Q_i(m+1)$, $(i = 1, 2)$, have within the limits of error the same values in the bands 3064 and 2811, and again in the bands 3122 and 2875. The two bands in each pair have therefore the same final state, and the equality of the differences $R_i(m) - P_i(m)$ in the bands 2811 and 3122 indicates a common initial state.

5. *New Experiments on the Water-Vapour Spectrum.*

To provide a check on these relations by comparing the differences of the wave-numbers, ν_0, of the null lines for pair of bands having the same initial

* 'Astrophys. J.,' vol. 3, p. 145 (1924).

† 'K. Akad. Wetensch. Amsterdam,' vol. 28, p. 174 (1925).

or the same final state, further data were necessary. The author therefore commenced an investigation on other bands of the system. The rapid falling off in intensity between the bands 3064 and 2811 indicated that the next band, λ 2608, discovered by Deslandres* and referred to by Liveing and Dewar,† would be so weak as to call for unduly long exposures with a grating spectrograph giving high dispersion. A preliminary investigation was made with a small Hilger quartz spectrograph (E 3), but the dispersion which amounted in the region of the band 2608 to about 9·6 A. per mm. did not resolve sufficiently some of the more closely packed lines to permit of satisfactory measurements being made. The small instrument was replaced by a large Hilger quartz spectrograph (E 1), which gives a much higher dispersion; the simplicity of the optical system permits of reasonably short exposures. The instrument is of the Littrow type and may be adjusted so that any desired region of the spectrum may be photographed. With the adjustments used in these experiments the range of wave-lengths included on a 10-inch plate was about 3390 to 2435 A.U., and the dispersion in the region occupied by the band 2608 was about 3·4 A. per mm. The arrangements for exciting the spectrum were very similar to those employed by Watson. A bulb containing air-free distilled water was connected by a capillary tube to one end of an end-on discharge tube provided with a quartz window. The other end was connected through drying tubes to a two-stage rotary oil pump. The pump was kept running during an exposure to prevent an accumulation of hydrogen in the discharge tube, and the dimensions of the capillary tube were so chosen as to give a suitable pressure of water vapour in the tube. It was found that a capillary of about 1 mm. bore and length 10 cm. was sufficient to prevent the pressure from becoming too high. The water vapour was excited by means of the discharge from an induction coil. Reasonably intense photographs of the band 2608 were obtained in about two hours. The iron-arc spectrum was used for comparison, a photograph of the iron arc being taken before and after that of water vapour so that any shift due to temperature variations might be detected. The wave-lengths of the iron lines used as standards were those given by Kayser and by Burns. A considerable number of very faint lines which appeared over the range included on the plates was attributed to the presence of stop-cock grease.

* 'C. R.,' vol. 100, p. 854 (1885).

† 'Phil. Trans.,' A, vol. 179, p. 27 (1888).

6. *Measurements of the Band* λ 2608.

New measurements of this band were made by the author from five different photographs and the wave-lengths obtained are probably accurate to within a few hundredths of an Angstrom unit, except perhaps for some of the fainter lines and lines which overlap. Besides the band 2608 a series of single lines near the head of the band 2811 was measured from three different photographs. Both the band and the singlet series are shown in Plate 6, No. 3. Plate 6 is a reproduction from photographs taken with the large quartz spectrograph. The figures indicate wave-lengths in Ångström units. The author's arrangement of the lines of the band 2608 into series is shown in Table I, where the wave-lengths in air and the vacuum wave-numbers, together with their first and second differences, are given. The intensities, I, are on the scale 1—10, 10 denoting the strongest lines, while 0 is used to indicate very faint lines. The letter *n* is

Table I.—Band λ 2608.

m.	$P_1(m)$.					$P_2(m)$.				
	I.	λ(air).	ν(vac).	Δ.	Δ^2.	I.	λ(air).	ν(vac).	Δ.	Δ^2.
2	10*n*	2624·06	38097·56			5	2615·48	38222·50		
				−31·90					−52·95	
3	8	2626·26	38065·66		−8·63	9	2619·11	38169·55		−4·22
				40·53					57·17	
4	9	2629·06	38025·13		8·57	8	2623·04	38112·38		2·58
				49·10					59·75	
5	9	2632·46	37976·03		8·79	9	2627·28	38052·63		8·64
				57·89					68·39	
6	9	2636·41	37919·14		6·84	9	2631·89	37984·24		3·46
				64·73					71·85	
7	7	2640·91	37854·41		7·35	9	2636·87	37912·39		5·75
				72·08					77·60	
8	7	2645·95	37782·33		6·56	8	2642·28	37834·79		6·11
				78·64					83·71	
9	7	2651·47	37703·69		6·46	9	2648·14	37751·08		4·86
				85·10					88·57	
10	8*n*	2657·47	37618·59		7·22	7	2654·37	37662·51		6·09
				92·32					94·66	
11	6	2664·01	37526·27		5·86	6	2661·06	37567·85		6·25
				98·18					100·91	
12	5	2671·00	37428·09		6·81	5	2668·23	37466·94		5·90
				104·99					106·81	
13	4	2678·51	37323·10		6·45	4	2675·86	37360·13		6·73
				111·44					113·54	
14	3	2686·53	37211·66		5·89	3	2684·01	37246·59		4·99
				117·33					118·53	
15	1	2695·03	37094·33		6·95	0	2692·58	37128·06		6·89
				123·98					125·42	
16	1	2704·07	36970·35		6·49	1	2701·71	37002·64		6·37
				130·47					131·79	
17	0	2713·65	36839·88			0	2711·37	36870·85		

Table I (continued).

m.	Q_1 (m).					Q_2 (m).				
	I.	$\lambda_{(air)}$.	$\nu_{(vac)}$.	Δ.	Δ^2.	I.	$\lambda_{(air)}$.	$\nu_{(vac)}$.	Δ.	Δ^2.
2	9	2622·06	38126·62			8	2613·41	38252·77		
				0·00					−22·82	
3	9	2622·06	38126·62		−10·04	10	2614·97	38229·95		−3·76
				−10·04					26·58	
4	8	2622·75	38116·58		8·98	9	2616·79	38203·37		4·92
				19·02					31·50	
5	10n	2624·06	38097·56		7·23	10	2618·95	38171·87		5·48
				26·25					36·98	
6	9	2625·93	38071·31		9·67	9	2621·49	38134·89		5·58
				35·92					42·56	
7	9	2628·35	38035·39		6·71	10	2624·42	38092·33		5·99
				42·63					48·55	
8	9	2631·30	37992·76		7·25	9n	2627·77	38043·78		6·22
				49·88					54·77	
9	9	2634·76	37942·88		7·18	9	2631·56	37989·01		5·89
				57·06					60·66	
10	9	2638·72	37885·82		7·00	8	2635·77	37928·35		6·83
				64·06					67·49	
11	8	2643·19	37821·76		6·62	6	2640·46	37860·86		6·48
				70·68					73·97	
12	9	2648·14	37751·08		7·53	6	2646·42	37786·89		6·67
				78·21					80·64	
13	6	2653·64	37672·87		7·46	6	2651·29	37706·25		7·02
				85·67					87·66	
14	3	2659·69	37587·20		6·64	8n	2657·47	37618·59		4·66
				92·31					92·32	
15	3n	2666·24	37494·89		6·55	6	2664·01	37526·27		8·80
				98·86					101·12	
16	1	2673·29	37396·03		8·36	1	2671·21	37425·15		7·39
				107·22					108·51	
17	1	2680·97	37288·81		6·18	0	2678·98	37316·64		7·96
				113·40					116·47	
18	0	2689·15	37175·41			0	2687·36	37200·17		

Table I—(continued.)

m.	$R_1(m)$.					$R_2(m)$.				
	I.	λ(air).	ν(vac).	Δ.	Δ^2.	I.	λ(air).	ν(vac).	Δ.	Δ^2.
2	1*n*	2617·88	38187·46			6	2609·23	38314·03		
				+30·08					+8·07	
3	3	2615·82	38217·54		−9·78	7	2608·68	38322·10		−5·29
				20·30					2·78	
4	4	2614·43	38237·84		9·02	8	2608·49	38324·88		5·56
				11·28					−2·78	
5	5	2613·66	38249·12		7·63	7	2608·68	37322·10		1·48
				3·65					4·26	
6	8	2613·41	38252·77		9·66	6	2608·97	38317·84		8·36
				−6·01					12·62	
7	5	2613·82	38246·76		7·15	7	2609·83	38305·22		6·00
				13·16					18·62	
8	*	—	(38233·60)		7·13	4	2611·10	38286·60		6·86
				20·29					25·48	
9	3	2616·11	38213·31		8·03	3	2612·84	38261·12		5·69
				28·32					31·17	
10	1	2618·05	38184·99		7·37	10	2614·97	38229·95		7·81
				35·69					38·98	
11	2	2620·50	38149·30		7·58	1	2617·64	38190·97		6·47
				43·27					45·45	
12	0	2623·47	38106·13		6·12	1	2620·76	38145·52		7·84
				49·39					53·19	
13	1	2626·96	37055·52			10	2624·42	38092·33		

* Not observed.

added when the line appears diffuse or where overlapping is suspected. Fig. 2 is the result of plotting m and ν.

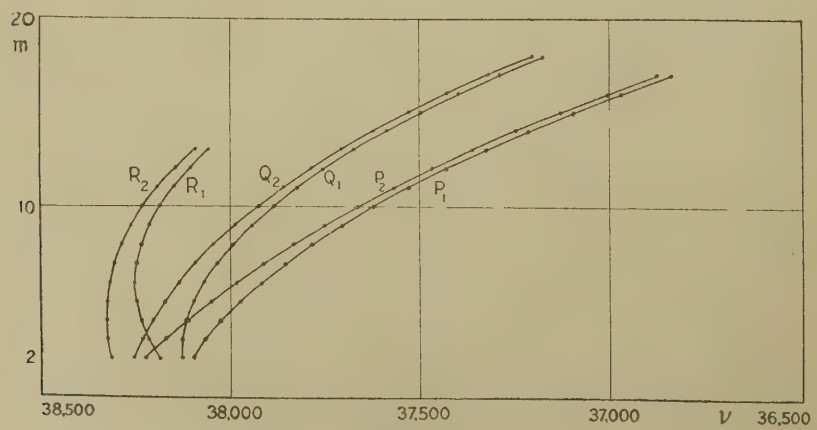


Fig. 2.

7. *Application of the Combination Principle.*

Adopting the notation of Richardson and Tanaka, the quantities P (m), Q (m), and R (m), which represent the wave-numbers of the band lines, are defined by the expressions,

$$\left.\begin{aligned} \mathrm{P}(m) &= \nu_0 + \mathrm{F}(m-1) - f(m), \quad & m-1 \to m, \\ \mathrm{Q}(m) &= \nu_0 + \mathrm{F}(m) - f(m), & m \to m, \\ \mathrm{R}(m) &= \nu_0 + \mathrm{F}(m+1) - f(m), & m+1 \to m, \end{aligned}\right\}, \tag{1}$$

where F represents the initial and f the final rotation term. If I and I′ denote the initial and final values of the moment of inertia of the emitter, and $\mathrm{B} = h/8\pi^2\mathrm{I}c$, $\mathrm{B}' = h/8\pi^2\mathrm{I}'c$, c being the velocity of light, then,

$$\left.\begin{aligned} \mathrm{P}(m) &= \nu_0 + (\mathrm{B}-\mathrm{B}')\,m^2 - 2\mathrm{B}m + \mathrm{B}, \\ \mathrm{Q}(m) &= \nu_0 + (\mathrm{B}-\mathrm{B}')\,m^2, \\ \mathrm{R}(m) &= \nu_0 + (\mathrm{B}-\mathrm{B}')\,m^2 + 2\mathrm{B}m + \mathrm{B}. \end{aligned}\right\}. \tag{2}$$

Equations (1) lead to the combination relation,

$$\mathrm{Q}(m-1) - \mathrm{P}(m) = \mathrm{R}(m-1) - \mathrm{Q}(m). \tag{3}$$

Since the mean values of the doublets fit more closely into parabolic formulæ than do the individual wave-numbers, the quantities P (m), etc., will be taken to represent $\frac{1}{2}\{\mathrm{P}_1(m) + \mathrm{P}_2(m)\}$, etc., in dealing with the general structure of the water-vapour bands. The differences Q ($m-1$) — P (m), R ($m-1$) — Q (m), and Q (m) — P (m) are given in Table II.

Table II.

m.	λ 3064	λ 2811	λ 2608	λ 3122	λ 2875
			Q ($m-1$) — P (m).		
3	72·48	72·27	72·09	—	69·77
4	109·63	109·65	109·53	105·57	106·12
5	146·79	146·54	145·65	141·26	140·78
6	183·27	183·30	183·03	176·45	176·39
7	219·44	219·60	219·70	211·21	210·74
8	255·25	255·33	255·30	245·51	245·63
9	290·56	290·70	290·88	278·97	279·40
10	325·48	325·73	325·40	313·25	312·45
11	359·79	359·94	360·03	346·07	345·93
12	393·81	393·88	393·79	378·63	—
13	427·11	427·00	427·37	410·64	—
14	459·90	459·73	460·43	442·18	—
15	492·09	492·35	491·70	473·02	—
16	523·54	524·05	524·08	503·02	—
17	554·31	554·39	555·22	582·64	—

Table II (continued).

m.	λ 3064	λ 2811	λ 2608	λ 3122	λ 2875
			R $(m-1)$ — Q (m).		
3	72·34	72·45	72·46	—	—
4	109·81	109·73	109·84	—	—
5	147·13	146·95	146·64	—	137·36
6	184·20	184·17	182·51	—	174·37
7	221·11	221·04	221·45	210·26	212·43
8	257·74	258·14	257·72	246·88	248·33
9	294·15	294·05	294·15	282·02	283·19
10	330·29	330·38	330·13	317·19	—
11	366·05	365·94	366·16	352·25	—
12	401·38	401·54	401·15	386·38	—
13	436·58	436·58	436·27	421·54	—
14	470·69	471·16	471·03	—	—
15	504·65	504·57	—	—	—
16	538·12	538·34	—	—	—
17	570·76	571·02	—	—	—
			Q (m) — P (m).		
2	33·92	32·05	29·67	—	30·52
3	67·64	64·47	60·68	64·13	61·07
4	101·39	96·27	91·22	96·91	91·41
5	134·94	128·13	120·39	128·26	122·68
6	168·01	159·68	151·41	160·90	151·17
7	200·75	190·87	180·46	191·05	180·87
8	233·15	221·62	209·71	221·58	209·89
9	265·11	252·01	238·56	251·75	237·97
10	296·62	281·66	266·54	282·85	266·45
11	327·60	311·12	294·25	311·25	295·98
12	358·14	339·92	321·47	340·11	—
13	388·01	367·97	347·94	368·56	—
14	417·33	395·48	373·77	396·17	—
15	446·01	422·96	399·38	422·92	—
16	473·71	449·42	424·09	449·39	—
17	501·12	474·50	447·36	474·83	—

There is reasonably good agreement between the differences in equation (3) for the first few values of m, but as m increases a marked deviation from equality makes its appearance. This is seen by comparing corresponding columns in the first two sections of Table II. Dieke considers the failure of the combination principle to be due to the Q series having final terms which differ from those in the P and R series.

8. *Moments of Inertia.*

The values of the moments of inertia of the emitter can be calculated from the figures in Table II, as it is obvious from equations (2) that

$$\mathrm{Q}(m-1) - \mathrm{P}(m) = \mathrm{R}(m-1) - \mathrm{Q}(m) \quad = \mathrm{B}'(2m-1)$$
$$\mathrm{Q}(m) \quad - \mathrm{P}(m) = \mathrm{R}(m-1) - \mathrm{Q}(m-1) = \mathrm{B}(2m-1).$$

Since the wave-number differences in these equations are not really equal, different values of B and B′ are obtained according as they are calculated from the P and Q differences or from the R and Q differences, but the difference in the actual values is not great, and the same relative values for the various bands result. The P and Q differences have been chosen here as they provide a greater quantity of material. The choice of the P and R differences is perhaps to be preferred if, as Dieke suggests, the final Q terms differ from the P and R terms, but actually very similar results are obtained and the data on the R branches in some of the bands are very limited. It will be observed that the values of B and B′, given in Table III, increase rapidly at first as m increases and gradually reach stationary values, after which they begin to decrease.

Table III.

m.	$h/8\pi^2 Ic$.					$h/8\pi^2 I'c$.				
	3064	2811	2608	3122	2875	3064	2811	2608	3122	2875
2	11·3	10·7	9·9	—	10·2	—	—	—	—	—
3	13·5	12·9	12·1	12·8	12·2	14·5	14·5	14·4	—	14·0
4	14·5	13·7	13·0	13·8	13·1	15·7	15·7	15·7	15·1	15·2
5	15·0	14·2	13·5	14·3	13·6	16·3	16·3	16·3	15·7	15·6
6	15·3	14·5	13·8	14·6	13·7	16·7	16·7	16·7	16·0	16·0
7	15·4	14·7	13·9	14·7	13·9	16·9	16·9	16·9	16·2	16·2
8	15·5	14·8	14·0	14·8	14·0	17·0	17·0	17·0	16·4	16·4
9	15·6	14·8	14·0	14·8	14·0	17·1	17·1	17·1	16·4	16·4
10	15·6	14·8	14·0	14·9	14·0	17·1	17·1	17·1	16·5	16·4
11	15·6	14·8	14·0	14·8	14·1	17·1	17·1	17·1	16·5	16·5
12	15·6	14·8	14·0	14·8	—	17·1	17·1	17·1	16·5	—
13	15·5	14·7	13·9	14·7	—	17·1	17·1	17·1	16·4	—
14	15·5	14·6	13·8	14·7	—	17·0	17·0	17·1	16·4	—
15	15·4	14·6	13·8	14·6	—	17·0	17·0	17·0	16·3	—
16	15·3	14·5	13·7	14·5	—	16·9	16·9	16·9	16·2	—
17	15·2	14·4	13·6	14·4	—	16·8	16·8	16·8	16·1	—

It is evident from the table that for any definite value of m the three bands 3064, 2811 and 2608 give the same final value for the moment of inertia. The same is true for the two bands 3122 and 2875, while the two bands 2811 and 3122 give the same initial value, and again the two bands 2608 and 2875. For the purpose of comparing the moments of inertia it is therefore immaterial what value of m is chosen. Since the turning points in B and B′ occur at the same place in all the columns, m has been chosen in this region. By choosing $m = 9$, small irregular variations are avoided. The results obtained are given in Table IV.

Table IV.—Moments of Inertia.

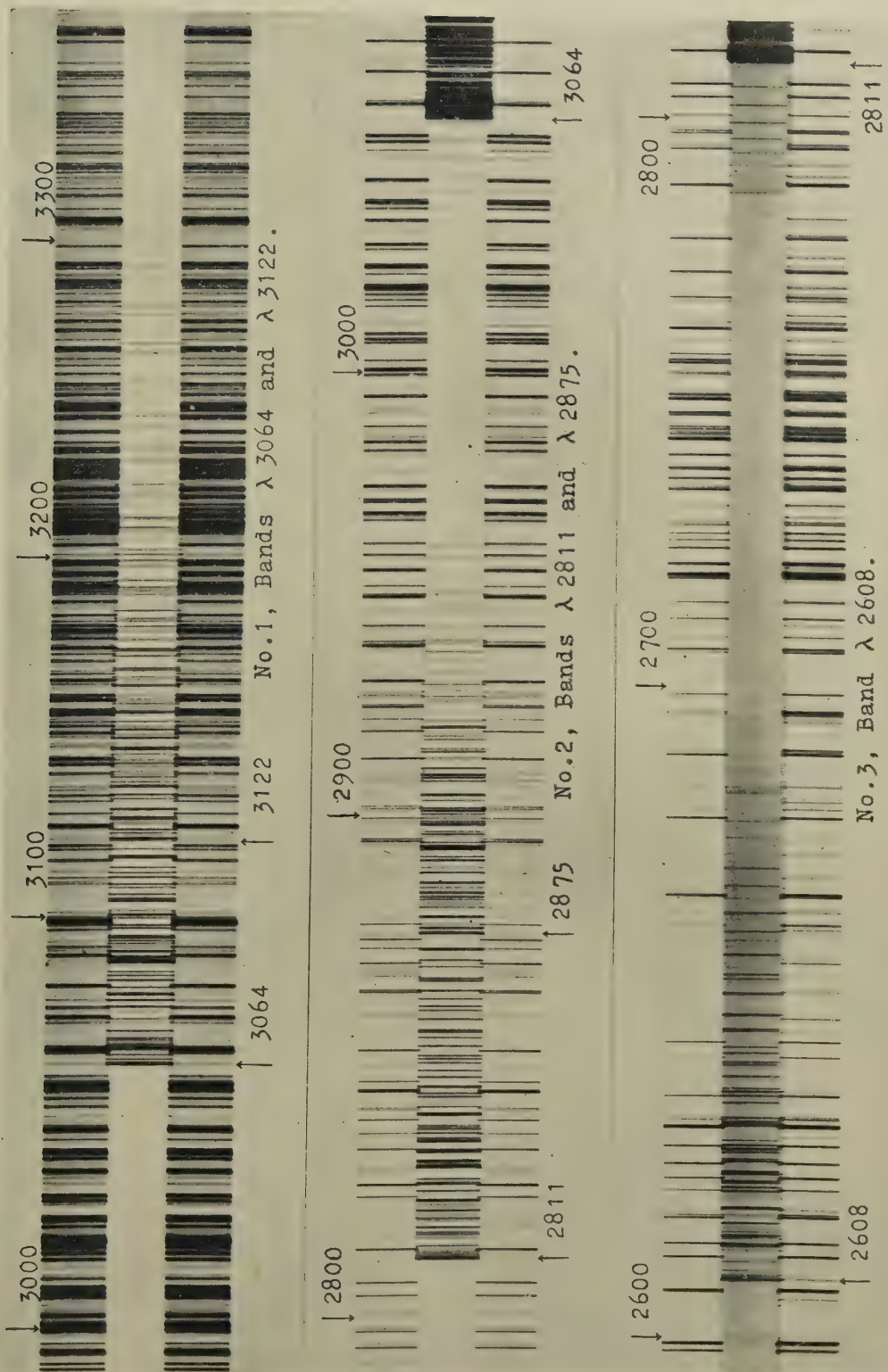
Band.	I.	I′.
3064	$1{\cdot}77 \times 10^{-40}$ gm. cm.^{2}	$1{\cdot}62 \times 10^{-40}$ gm. cm.^{2}
2811	1·87	1·62
2608	1·98	1·62
3122	1·87	1·69
2875	1·98	1·69

The moments of inertia for the bands 3064 and 2811 do not differ appreciably from those given by Watson. Watson* calculated the dimensions of a linear H_2O molecule and of an OH ion which would have the required moment of inertia and found for the distance between the O and H nuclei $0{\cdot}7 \times 10^{-8}$ cm. in the H_2O molecule and $1{\cdot}02 \times 10^{-8}$ cm. in the OH ion. The distance between the O and H nuclei in the ice crystal is, from the measurements of W. H. Bragg, $1{\cdot}38 \times 10^{-8}$ cm. In both cases the size is less than would be expected when the low gaseous pressure and the effect of rotation are taken into account, but the evidence is distinctly in favour of the OH ion as emitter rather than the H_2O molecule.

9. *Band System.*

The results of Table IV provide an extension of the scheme suggested by Dieke, and the addition of the band 2608 allows of the scheme being verified by taking differences of the wave-numbers, ν_0, of the null lines in the various bands. Since the wave-numbers of the lines do not fit quite accurately into the formulæ, some doubt must be attached to the values of ν_0, obtained by extrapolation, but this uncertainty may be avoided by taking corresponding lines in the various bands instead of the actual null lines. If the values of $Q(m)$ are taken, it can easily be shown that the results will be similar to those obtained by using the null lines. Consequently the wave-numbers, $Q_2(2)$, are arranged in Table V so that those corresponding to bands with the same initial moment of inertia of the emitter appear in the same row, and those from bands with the same final value in the same column. For convenience of reference the wave-lengths of the band heads are given in brackets alongside the wave-numbers, $Q_2(2)$.

* ‘Astrophys. Journ.,’ vol. 3, p. 145 (1924).


3000
3100
3200
3300
3064
3122
No.1, Bands λ 3064 and λ 3122.
2800
2900
3000
2811
2875
3064
No.2, Bands λ 2811 and λ 2875.
2600
2700
2800
2608
2811
No.3, Band λ 2608.

(*Facing p.* 384.)

Table V.—Band System.

(λ 3064),	32474·70			
	2986·85			
(λ 2811),	35461·55	*3568·30*	31893·25,	(λ3122)
	2791·22		*2791·01*	
(λ 2608),	38252·77	*3568·51*	34684·26,	(λ2875).

The differences between neighbouring numbers in the two columns are in close agreement, as also are the differences between the numbers in rows 2 and 3. These results provide a verification of the arrangement of the band system. The superposition in the photographs of the bands 3122 and 2875 on the bands 3064 and 2811 is merely accidental, and the above scheme shows the way in which the various bands are actually related. The scheme suggests the probability of the occurrence of another band having the same initial state as the band 3064, and makes it possible to predict with considerable accuracy the position of the lines of this band. A band appears at λ 3428, which is in the region indicated by the scheme. In an investigation on impurity lines in connection with his work on the many-lined spectrum of hydrogen, Tanaka* has made some measurements of the band 3428 and gives the wave-lengths of 45 lines. The predicted values of the first 10 lines of Q_2 series agree very closely with lines observed by Tanaka, and there is also good agreement between the predicted and observed values for the lines from $R_2(6)$ to $R_2(12)$. There is evidently a considerable amount of overlapping, especially in the region $Q_1(3)$ to $Q_1(6)$. These lines fall in the neighbourhood of a strong diffuse line in Tanaka's measurements. The author has accounted for nearly all Tanaka's lines, but owing to the overlapping some uncertainty must be attached to the complete arrangement of the band, which is therefore not given here. It seems definitely established, however, that this band falls into line with the scheme.

10. *General Discussion—Numeration.*

Although the internal structure of the bands does not conform to the simple theory, it appears that the relations between the different bands are in good agreement. In fig. 1 the mean values of the wave-numbers of the doublet components have been plotted, and are indicated by the broken lines. The points of intersection of these curves are approximately, at $m = 1$ for the P and Q branches, at $m = \frac{1}{2}$ for P and R, and at $m = 0$ for Q and R. The theoretical curves of equations (2) intersect at $m = +\frac{1}{2}, 0$ and $-\frac{1}{2}$ instead of

* 'Roy. Soc. Proc.,' A, vol. 108, p. 594 (1925).

1, $\frac{1}{2}$, 0. It is evident, then, that a better fit between the two sets of curves is obtained if m, in fig. 1, is replaced by $m' = m - \frac{1}{2}$. If, now, the values of B and B′ are calculated with this change in the values of m, it is found that the results show much less variation than is exhibited in Table III. For example, col. 1 shows a variation of 4·3 between the maximum and minimum values, whereas the results obtained by using m' show a variation of only 1·3. Moreover, these results vary continuously instead of having a turning value. and while the results in the table represent a rapid decrease in the moment of inertia for the first few increasing values of m, those obtained from m' increase continuously with m'. It is reasonable to expect an increase rather than a decrease in the moment of inertia with increasing energy of rotation. If the simple formula $F(m) = Bm^2$ be replaced by the expression of Kramers and Pauli,*

$$F(m) = B[\sqrt{(m^2 - \sigma^2)} - \rho]^2,$$

the expression in the square brackets must be used instead of m in calculating the moments of inertia, and it is highly probable that this quantity,

$$[\sqrt{(m^2 - \sigma^2)} - \rho]^2,$$

may have values differing from integers. In fact, as Mulliken† points out, successive values of this expression may not even differ by unity. The quantities $\sigma h/2\pi$ and $\rho h/2\pi$ represent the components of the electron momentum in directions at right angles to, and parallel to, the momentum $mh/2\pi$. Sandeman‡ points out that the difference between successive values of $F(m) - F(m-1)$, that is, of $Q(m) - P(m)$, should approach asymptotically the value 2B for higher values of m. Actually these differences show no indication of tending to a steady value.

11. *Doublets.*

So far only the mean values of the doublets have been considered, as those serve to illustrate the general structure of the bands equally as well as the individual lines. These mean values cannot, however, be taken to have any physical meaning, as the absence of the conditions which give rise to the doublets would lead to a line probably occupying a position other than that obtained by taking the mean wave-number of the two components. If, as Kratzer§ assumes, the doublets may be explained by giving values $\pm\varepsilon$ to ρ

* ‘Z. f. Physik,’ vol. 13, p. 351 (1923).
† ‘Phys. Rev.,’ vol. 28, pp. 481 and 1202 (1926).
‡ ‘Roy. Soc. Proc.,’ A, vol. 108, p. 607 (1925).
§ ‘Ann. d. Physik,’ vol. 71, p. 72 (1923).

in the expression of Kramers and Pauli, this would certainly be the case, especially if ε were large. Difficulty is met with in applying this theory to the water-vapour bands on account of the great variation of the doublet separation with m. Mulliken states that the values given to σ and ε must vary with m in these bands. The doublet separations in the five bands are given in Table VI.

Reference has already been made to the empirical relations given by Heurlinger for the doublet separations in the band 3064. The most striking feature exhibited in the table is the very close agreement between corresponding

Table VI.—Doublet Separations.

m.	λ 3064.	λ 2811.	λ 2608.	3122.	λ 2875.
			$P_2(m) - P_1(m)$.		
2	126·66	126·73	124·94	—	127·00
3	104·19	104·34	103·89	105·59	105·24
4	87·52	87·45	86·45	88·40	88·88
5	75·11	75·01	76·60	76·36	75·64
6	65·56	65·40	65·10	66·76	66·81
7	58·15	58·14	57·68	59·63	59·09
8	52·80	52·39	52·46	53·86	53·75
9	47·88	47·93	47·39	49·39	48·96
10	44·38	44·25	43·92	45·25	45·16
11	41·62	41·34	41·58	42·20	42·25
12	39·04	38·93	38·85	39·77	—
13	36·99	36·05	37·03	37·56	—
14	35·28	35·25	34·93	35·72	—
15	33·80	33·65	33·73	34·18	—
16	32·60	32·95	32·29	33·03	—
17	31·74	31·38	30·97	32·17	—
			$Q_2(m) - Q_1(m)$.		
2	126·44	126·84	126·15	—	126·86
3	103·81	103·94	103·33	105·45	104·78
4	87·34	87·12	86·79	88·13	87·04
5	74·48	74·25	74·31	75·78	75·50
6	64·66	64·42	63·58	65·88	65·21
7	57·04	57·04	56·94	58·29	57·54
8	51·07	50·98	51·02	52·29	51·84
9	46·27	45·94	46·13	47·83	47·75
10	42·44	42·20	42·53	43·38	44·24
11	39·16	39·18	39·10	39·89	43·59
12	36·52	36·48	35·81	37·15	—
13	34·25	34·18	33·38	35·10	—
14	32·42	32·25	31·39	33·18	—
15	30·75	30·64	31·38	31·09	—
16	29·49	29·27	29·12	29·87	—
17	28·12	28·20	27·83	28·66	—
18	27·09	26·75	25·76	27·77	—

Table VI—(continued).

m.	λ 3064.	λ 2811.	λ 2608.	λ 3122.	λ 2875.
			$R_2(m) - R_1(m)$.		
2	126·61	126·81	126·57	—	—
3	104·73	104·07	104·56	—	—
4	88·16	87·87	87·04	—	81·19
5	75·50	75·44	72·98	—	73·83
6	66·10	65·33	65·07	—	66·85
7	58·73	58·74	58·46	—	61·13
8	52·90	53·17	(53·00)	—	54·04
9	48·42	48·49	47·81	—	—
10	44·81	44·92	44·96	—	—
11	41·95	41·72	41·67	—	—
12	39·68	39·78	39·39	—	—
13	37·42	36·83	36·71	—	—

doublets in the three bands 3064, 2811 and 2608, and between those in the two bands 3122 and 2875, although in these latter bands there are greater irregular variations. The separations in the two groups differ only slightly, but a real difference seems to exist. It is interesting to note that this grouping is exactly that obtained by arranging together bands with the same final moment of inertia, a fact which indicates that probably the factors governing the separation of the doublets are to be looked for in the final state rather than in the initial state. The combination principle has been shown to hold approximately for low values of m, and here also there is fairly close agreement between the separation of the doublets in the Q branch with those in the P and R branches. The deviation from the combination principle increases with m, and there is a corresponding increase in the differences between the Q and P (or R) doublet separations. It is probable, then, that the difference in the doublet separations arises from the same cause as does the failure of the combination principle, and that these may both be accounted for by Dieke's assumption that the Q lines have final terms which differ from those for the P and R lines. The introduction of quantities σ and ε varying with m may easily cause this difference in the final Q terms, since in the P and R branches m suffers transitions ± 1, whereas in the Q branch no transition of m takes place. It is difficult to see how such definite relations can exist between σ and ε and m as would be necessary to afford an explanation of the doublets, and until these relations are understood the fine structure of the water-vapour bands cannot be considered as completely explained.

12. *Singlet Series.*

The author's measurements of the series of single lines, already referred to, are given in Table VII. These lines are found to fit reasonably well into a parabolic formula. The second differences, however, show a distinct tendency to increase numerically in passing along the series. Exactly the same kind of variation of the second differences appears in the series given by Watson. The strong similarity between the two series points to a common origin, and the relation of their intensities to those in the water-vapour bands indicates the probability that these series also are in some way due to water vapour.

Table VII.

I.	$\lambda_{(air)}$.	$\nu_{(vac)}$.	Δ.	Δ^2.
10	2806·45	35621·67		
			+44·98	
10	2802·92	35666·65		−3·34
			41·64	
10	2799·65	35708·29		3·34
			38·30	
10	2796·65	35746·59		4·41
			33·89	
9	2794·00	35780·48		4·55
			29·34	
8	2791·71	35809·82		4·96
			24·38	
7	2789·81	35834·20		4·98
			19·40	
6	2788·30	35853·60		5·64
			13·76	
5	2787·23	35867·36		7·84
			5·92	
4*n*	2786·77	35873·28		2·57
			3·35	
4*n*	2786·51	35876·63		6·70
			−3·35	
4*n*	2786·77	35873·28		7·33
			10·68	
3	2787·60	35862·60		5·51
			16·19	
2*n*	2788·86	35846·41		7·06
			23·25	
1*n*	2790·67	35823·16		6·75
			30·00	
1*n*	2793·01	35793·16		6·60
			36·60	
0*n*	2795·87	35756·56		

Summary.

Evidence on the nature of the emitter of the water-vapour bands is discussed and shown to be in favour of the OH ion. An account is given of the work of Heurlinger, Watson and others on the bands 3064, 3122, 2811 and 2875.

Further work on the water-vapour spectrum by the author is described, and details are given of the band 2608 (Table I). This band is similar in structure to the others, and it is found to lead to the same final amount of inertia as the bands 3064 and 2811. The initial value is the same as that for the band 2875. The scheme of bands suggested by Dieke has been extended and verified by taking differences of the wave-numbers of corresponding lines in the various bands (Table V). It has been shown that the band 3428, of which some measurements have been given by Tanaka, also fits into the scheme.

The doublet separations are found to be the same for bands which give the same final value for the moment of inertia. The possibility of explaining the formation of the doublets in terms of the Kratzer-Kramers and Pauli expression is discussed.

The author has also measured a group of single lines near the head of the band 2811 (Table VII), which fits into a parabolic formula and bears a strong resemblance to the series of single lines measured by Watson.

The paper is illustrated by photographs and curves.

The author wishes to express his thanks to Prof. H. Stanley Allen for suggesting the work of this paper and for his kindly interest and many valuable suggestions throughout the course of the investigations, and to Prof. O. W. Richardson for his helpful criticism. The author also wishes to thank Dr. Ian Sandeman for his assistance and advice in connection with some preliminary experiments.

The Brownian Movement of a Galvanometer Coil and the Influence of the Temperature of the Outer Circuit.

By L. S. Ornstein, H. C. Burger, J. Taylor and W. Clarkson.

(Physical Institute of the University of Utrecht.)

(Communicated by O. W. Richardson, F.R.S.—Received March 25, 1927.)

Introduction.

Some time ago, in connection with certain curves published from this Institute, and illustrative of the thermo-relay method for the magnification of galvanometer deflections, Ising pointed out* that the current irregularities superposed upon the baseline-current, as traced out photographically, could be interpreted as due to the Brownian movement of the suspended coil of the galvanometer, and indeed were of the order required theoretically.†

To prove definitely that these small variations are really due to the Brownian movements, it is evidently necessary to show that the temperature of the galvanometer system has an influence upon the mean value of the movement. The difficulties involved in maintaining the galvanometer itself at sufficiently low temperatures are great, and it is much more convenient to cool a series-connected coil in the outer circuit. This introduces the problem of the Brownian movement in a system at two different temperatures, and this has not, until recently, been investigated theoretically. The purpose of the present work is, then, to develop the necessary theoretical methods for such problems, and to obtain an expression for the Brownian movement which can be subjected to direct experimental verification.

Theoretical Considerations.

In order to find the extension of the known theory of the Brownian movement, it will be good to consider first a variation of the classical theory of this phenomenon in a form which is suitable for extension to the more complicated case which is to be treated in the interpretation of our experiments.

Let us with Einstein‡ consider a colloidal particle of mass m. Let βu be the

* 'Phil. Mag.,' vol. 1, p. 827 (1926).

† *Cf.* plate in this paper.

‡ *Cf.* L. S. Ornstein, 'Z. f. Physik,' vol. 41, p. 848 (1927).

resistance which our particle experiences when it moves with a constant velocity u, in an ideal continuous fluid of given inner friction. When, in such a case, the particle is set free with a velocity u_0, the equation of motion is,

$$m\, du/dt = -\beta u,$$

and, therefore,

$$u = u_0 e^{-\beta t/m}.$$

When now the particle moves in a real liquid, the resistance against the motion is not quite the same, because of the fact that the fluid has a molecular structure which causes irregular impulses to be exercised on the particle. However, in the mean, the irregular forces can be resolved into two parts, one equal to the resistance in an ideal liquid of the same frictional coefficient, and one quite irregular $F(t)$, such that the mean value taken for a great number of particles, all starting with the velocity u, will be zero. We must therefore replace the equation of motion by

$$m\, du/dt = -\beta u + F, \tag{1}$$

where $\overline{F}$, the mean value of the irregular force $F(t)$ is zero. An equation of the same form can be applied to other cases. When we have, for example, a circuit with self-inductance L and resistance r (placed in a thermostat) the equation for the current i is

$$L\, di/dt = -ri,$$

if the molecular structure is neglected. When, however, we take this structure into account, small variations in the circuit will cause electromotive forces, and we have an equation of the form

$$L\, di/dt + ir = E, \tag{2}$$

$\overline{E} = 0$, where E is the mean electromotive force caused by the fluctuations.

Now it is easy to integrate (1) or (2). We obtain, for example, for (1).

$$u = u_0 e^{-\beta t/m} + 1/m \,.\, e^{-\beta t/m} \,.\, \int_0^t F(z) \,.\, e^{\beta z/m} \,.\, dz, \tag{3}$$

where u_0 is the velocity of the particle at the time 0, and t is the time at which the velocity is considered.

When we take the mean value we get, $\overline{F}$ being zero,

$$u = u_0 e^{-\beta t/m}.$$

The mean movement in a liquid composed of molecules is, therefore, the same as the real movement in an ideal viscous fluid. Now it is very interesting to calculate the mean energy of our particle, for if we know its value we can

deduce another property of the function F. Applying the theorem of equipartition which states that the mean value of the kinetic energy $\frac{1}{2}mu^2$ is equal to $\frac{1}{2}k\mathrm{T}$, where k is the Boltzmann constant and T the absolute temperature. We get from (3) :—

$$\overline{u^2} = u_0^2 \cdot e^{-2\beta t/m} + \frac{1}{m^2} \cdot e^{-2\beta t/m} \cdot \int_0^t \int_0^t \overline{\mathrm{F}(z)\,\mathrm{F}(y)} \cdot e^{\beta/m \cdot (z+y)} \cdot dz \cdot dy. \qquad (4)$$

The mean value $\overline{\mathrm{F}(z)\,\mathrm{F}(y)}$ must be different from zero, since u^2 for $t = \infty$ is positive, but it is clear that the correlation of the forces F (z) and F (y) can be appreciable only when the difference of z and y is small. If we put $y = z + \psi$, the integral in (4) can be transformed into

$$\int \overline{\mathrm{F}(z)\,\mathrm{F}(z+\psi)} \cdot d\psi \cdot \int_0^t e^{2\beta z/m} \cdot dz,$$

ψ being so small that in the exponent $z + y$ can be replaced by $2z$. It is easily seen that the integral $\int \mathrm{F}(z)\,\mathrm{F}(z+\psi)\,d\psi$ is independent of z and y, and is a characteristic constant for the problem which we shall denote by $\overline{\mathrm{FF}}$. After some simple reductions we now find :

$$\frac{m}{2}\overline{u^2} = \frac{m}{2}u_0^2 e^{-2\beta t/m} + \frac{1}{4\beta}(1 - e^{-2\beta t/m}) \cdot \overline{\mathrm{FF}}.$$

For times which are large compared with the time $m/2\beta$ we find

$$\frac{m}{2} \cdot \overline{u^2} = \frac{1}{4\beta} \cdot \overline{\mathrm{FF}},$$

which expression, by applying the theorem of equipartition, can be reduced to

$$\overline{\mathrm{FF}} = 2\,\beta\,k\,\mathrm{T}.$$

On the same lines we find for the electromotive force

$$\overline{\mathrm{EE}} = 2\,r\,k\,\mathrm{T}.$$

It is well to remark that the hypothesis about the quantity $\overline{\mathrm{FF}}$ deduced from the theorem of equipartition permits of being found the mean square of the displacement, a quantity which can be measured (such is not the case for the velocity u in the movement which we shall consider in connection with the galvanometer Brownian movement).* Let us now consider a somewhat more

* Putting $u = dx/dt$ we find by integration of (4) the well-known formula for the mean deviation.

complicated case, that of a galvanometer joined in series with a resistance, but let us assume that the whole system is at the same temperature. If, now, A is the moment of inertia, α the torsional constant, β the air-damping, κ the galvanometer coil effective flux, r the resistance in the circuit, L the value of the inductance in the external circuit, θ the angle of deflection, and i the current, we get:

$$A\ddot{\theta} + \beta\dot{\theta} + \alpha\theta + \kappa i = F,$$

$$L\, di/dt + ri - \kappa\dot{\theta} = E,$$

where F is the couple on the galvanometer coil caused by irregular impacts of the air molecules and E an electromotive force of the same species.

We shall show that if we assume $\overline{F} = \overline{E} = 0$, $\overline{FF} = 2\beta\kappa T$, and $\overline{EF} = 0$, we get for $A\overline{\theta^2}/2$ and $\alpha\overline{\theta^2}/2$ the equipartition value $kT/2$, and also $L\overline{i^2}/2 = kT/^2$.

We shall consider the aperiodic case where $4A\alpha = (\beta + \kappa^2/2)^2$, and put $\alpha/A = \nu^2$. Now a solution can be found taking $\theta = ae^{pt}$, $i = be^{pt}$. We find for p three values that, when the resistance r is considered to be large, take the form

$$p_1 = -\nu + \kappa L/r, \quad p_2 = -\nu - \kappa L/r, \quad \text{and} \quad p_3 = \frac{-r/L}{1 - \frac{\kappa^2 L^2}{r^2}},$$

while b is given by the equation $b = \frac{\kappa p}{Lp + r} \cdot p_3$ can be put equal to $-r/L$, only, in the expression for b, the second approximation must be considered.

The values of a_1, a_2, a_3, b_1, b_2, and b_3, for the case in which r and E are taken into account may be determined by the known method of the variation of parameters. When we neglect the initial values of θ, $\dot{\theta}$ and i, which will tend towards zero after small times, we get for the current

$$i = \frac{1}{L}\, e^{-rt/L} \int_0^t e^{rz/L} \cdot E(z) \cdot dz,$$

and for

$$\theta = \frac{1}{(p_1 - p_2)} \frac{1}{A} \left(e^{-p_1 t} \int_0^t e^{p_1 z} \left(F(z) - \frac{\kappa}{r} - E(z) \right) \right) \cdot dz$$

$$- e^{-p_2 t} \int_0^t e^{p_2 z} \left(F(z) - \frac{\kappa}{r} - E(z) \right) \cdot dz.$$

Members of the order L/r have been neglected.

If we take into consideration our hypothesis for $\overline{EE}$ we find,

$$\overline{i^2} = k\mathrm{T}/\mathrm{L},$$

and for

$$\overline{\theta^2} = \frac{1}{(p_1 - p_2)^2}\frac{1}{\mathrm{A}^2}\left(-\frac{1}{p_1} - \frac{1}{p_2} + \frac{r}{p_1 + p_2}\right)(\overline{FF} + \kappa^2\overline{EE})$$

$$= \frac{1}{2\nu^3\mathrm{A}^3}k\mathrm{T}\,(\beta + \kappa^2/r) = \frac{k\mathrm{T}}{\alpha},$$

or $\alpha\overline{\theta^2}/2 = k\mathrm{T}/2$. In an analogous way the equipartition value can be found for $\mathrm{A}\overline{\theta^2}/2$.

The given deduction furnishes a good example for the theorem that the mean value of the kinetic energy in any Brownian movement is always $k\mathrm{T}/2$ for each degree of freedom, whatever may be the number of causes for the irregular movement.

We can now proceed to the treatment of the problem which is of great importance for experiments of Brownian movement, the problem of the influence of two temperatures in the system. As an introduction we shall treat the problem of two coils in series, with self-inductances L_1 and L_2 respectively, and resistances r_1 and r_2 which are at temperatures T_1 and T_2. Now if the rest of the system was at the same temperature the accidental E_1 for the first coil would have the values $\overline{\mathrm{E}_1} = 0$, and $\overline{\mathrm{E}_1\mathrm{E}_1} = 2r_1k\mathrm{T}$, and for the second $\overline{\mathrm{E}_2} = 0$, and $\overline{\mathrm{E}_2\mathrm{E}_2} = 2r_2k\mathrm{T}$. Our hypothesis now will be that the same relations are still true when the different coils are at different temperatures. We shall assume that the mean value of $\overline{\mathrm{E}_1\mathrm{E}_2}$ is zero. The fact that an analogous hypothesis in our considerations of the galvanometer has given correct results has led to this assumption. Now the equation for motion for the electricity in our coils is

$$(\mathrm{L}_1 + \mathrm{L}_2)\,di/dt + (r_1 + r_2)i = \mathrm{E}_1 + \mathrm{E}_2,$$

where i is the current. Integrating the equation we find for

$$i = i_0 e^{\frac{-r_1+r_2}{\mathrm{L}_1+\mathrm{L}_2}\cdot t} + e^{-\frac{r_1+r_2}{\mathrm{L}_1+\mathrm{L}_2}\cdot t}\int_0^t e^{\frac{r_1+r_2}{\mathrm{L}_1+\mathrm{L}_2}\cdot z}(\mathrm{E}_1(z) + \mathrm{E}_2(z))\,dz,$$

and for $\overline{i^2}$ again neglecting the terms with i_0,

$$\overline{i^2} = \frac{e^{-\frac{2(r_1+r_2)}{\mathrm{L}_1+\mathrm{L}_2}\cdot t}}{(\mathrm{L}_1+\mathrm{L}_2)^2}\int_0^t\int_0^t e^{\frac{r_1+r_2}{\mathrm{L}_1+\mathrm{L}_2}(z+y)}\,\overline{(\mathrm{E}_1(z) + \mathrm{E}_2(z))\,((\mathrm{E}_1(x) - \mathrm{E}_2(x))}\,.\,dz\,.\,dx$$

$$= \frac{1}{\mathrm{L}_1+\mathrm{L}_2}\frac{1}{r_1+r_2}(2kr_1\mathrm{T}_1 + 2kr_2\mathrm{T}_2),$$

and for

$$\frac{(\mathrm{L}_1+\mathrm{L}_2)\,\overline{i^2}}{2} = \frac{k}{2}\frac{r_1\mathrm{T}_1 + r_2\mathrm{T}_2}{r_1 + r_2}$$

for $T_1 = T_2$, we find the equipartition value of the energy. However, when the temperatures are unequal we find a value intermediate between those corresponding to T_1 and T_2, the resistance of the coil playing the part of weight factors.

Now we are prepared to calculate the Brownian movement for the galvanometer. Let r_1 be the resistance of the galvanometer (temperature T_1) and r_2 that of the outer circuit (temperature T_2). Then the equation of motion of the galvanometer and the electrical motion is, neglecting the inductance, the following:

$$A\ddot{\theta} + \beta\dot{\theta} + \alpha\theta + \kappa i = E$$
$$(r_1 + r_2)\, i - \kappa\dot{\theta} = E_1 + E_2.$$

We get for the equation of the galvanometer by the elimination of i,

$$A\ddot{\theta} + \left(\beta + \frac{\kappa^2}{r_1 + r_2}\right)\dot{\theta} + \alpha\theta = F - \frac{\kappa}{r_1 + r_2}(E_1 + E_2),$$

if we assume

$$\overline{FF} = 2\beta kT, \qquad \overline{FE_1} = 0$$
$$\overline{E_1E_1} = 2r_1kT_1 \qquad \overline{FE_2} = 0$$
$$\overline{E_2E_2} = 2r_2kT_2 \qquad \overline{E_1E_2} = 0$$

we get for

$$\alpha\overline{\theta^2} = \frac{1}{\alpha\left(\beta + \frac{\kappa^2}{r_1 + r_2}\right)}\left(\overline{FF} + \frac{\kappa^2}{(r_1 + r_2)^2}(\overline{E_1E_1} + \overline{E_2E_2})\right)$$
$$= \frac{\kappa}{\beta + \frac{\kappa^2}{r_1 + r_2}}\left(\beta T_1 + \frac{\kappa^2}{(r_1 + r_2)^2}(r_1T_1 + r_2T_2)\right).$$

When the damping β is small compared with $(\kappa^2/r_1 + r_2)$, we have

$$\frac{\alpha\overline{\theta^2}}{2} = \frac{k}{2}\frac{r_1T_1 + r_2T_2}{r_1 + r_2}, \tag{5}$$

a relation which can be checked by experiment, as described later in this paper.

It is not difficult to apply analogous considerations to the Einthoven string galvanometer.

It is well to remark that at very low temperatures no quantum effects are probable, except those for the resistance itself, for the logical extrapolation of our result to low temperatures is that the mean value $\overline{EE}$ shall not be put $4r \cdot kT/2$, that is, four times the resistance multiplied by the equipartition value

of the energy at the temperature given, but should be replaced by the quantum value for that temperature. Now in the Einstein expression that has to be used, the frequency (that of the galvanometer) is so low that there is no difference between the Einstein energy function and the equipartition H value for all temperatures that can be attained.

The Experimental Method.

Principle of the method.

In order to obtain a sufficiently high magnification of the small Brownian movements of the galvanometer coil the thermo-relay method, previously described by Moll and Burger,* was utilised. Shortly, this method is as follows :—

A beam of light from an incandescent lamp illuminates a rectangular opening fixed at the surface of a convex lens which converges the beam upon the mirror of the galvanometer, of which the Brownian movement is being measured. The reflected light is converged by a second lens placed after and just in front of the galvanometer mirror, to give a real image of the rectangular slit exactly in the plane of, and on the surface of, the thermocouple strip of the relay. The latter consists of a very thin composite thermocouple strip ABCD (see fig. 1), AB and CD are of constantan and BC of manganin. The foil is blackened so as to absorb radiation and is mounted on a lamp bridge in an evacuated tube.

Linear concentration of the rectangular image is obtained by a cylindrical lens placed before the thermo-relay, so that the image is condensed upon the central portion of the strip. The relay is adjusted in its holder until no current is detectable in the galvanometer (second galvanometer) connected to it. This adjustment implies that both the junctions are equally heated so that their E.M.Fs. balance.

Any rotation of the first galvanometer mirror causes a displacement of the image on the relay strip and a non-symmetrical heating of the junctions, so that a current is indicated by the second galvanometer. The deflections are proportional to the rotation of the first galvanometer provided these are not too large.

Utilising this method, it is possible to obtain a very high magnification factor, depending upon the intensity of the light source and the sensitivity of the thermo-relay.

In the experiments to be described a relay of specially thin thermo-foil* was

* 'Phil. Mag.,' vol. 1, p. 624 (1925).

used, and with a 12-volt half-watt lamp as source a magnification of over 500 was easily obtainable. To assure constancy of the sensitivity it is, of

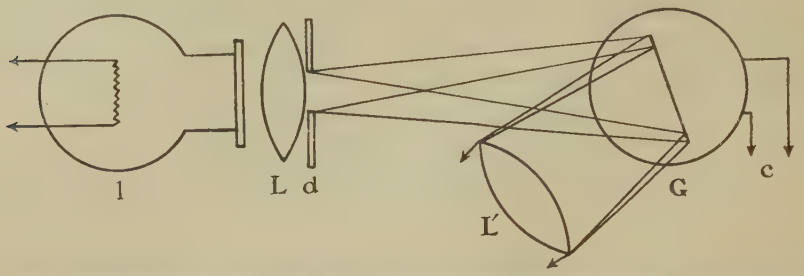


Fig. 1a.—Optical Arrangement for First Galvanometer. *l*, lamp with constant current: L, first lens; *d*, diaphragm; G, first galvanometer; *c*, to coil connections.

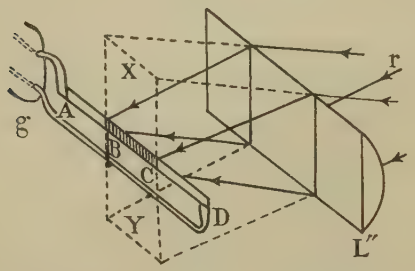


Fig. 1b.—Focussing System on Thermo-relay. *r*, rays from second lens; L″, cylindrical lens condensing image XY on to BC; XY, image of diaphragm, equal in width to BC; ABCD, thermo-element; *g*, leads to second galvanometer.

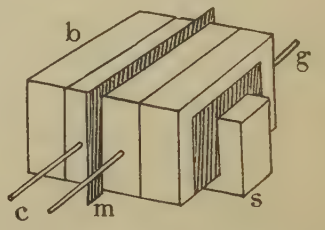


Fig. 1c.—Coils-Galvanometer Connections. *b*, plane brass blocks; *c*, to coils; *m*, mica; *s*, screw clamp; *g*, to first galvanometer.

course, necessary to have all parts of the lamp-relay system rigidly fixed in their relative positions and to have a light source of invariable intensity.

The deflections of the light spot of the second galvanometer were registered photographically on sensitised paper on a rotating drum, the "light-line" of the galvanometer being made punctiform on the drum by means of a cylindrical lens placed before it.

"*Galvanometer and Relay*" *System.*

The first galvanometer was a suspended coil galvanometer of special construction, designed by Moll,* and characterised by a high zero steadiness. Its resistance was 45 ohms.

The galvanometer was enclosed in a heavy cylindrical metal box covered with

* 'Proc. Phys. Soc.,' vol. 25, p. 253 (1923).

thick felt, and having a plane glass window through which the light beam could pass. The precaution of enclosing the galvanometer was necessary to avoid disturbances and E.M.Fs. set up by the heating effect of the air currents, at the terminals, &c.

This galvanometer and the auxiliary apparatus were fixed to a stout brass plate upon a very stable stone wall bracket.

Sensitivity of the Thermo-relay.

An auxiliary circuit, which could be connected to the first galvanometer circuit, comprising a resistance box, galvanometer shunt, and small E.M.F. source, served to measure the sensitivity of the system. The optical system requires careful adjustment and fixing, and, as noted above, the image of the diaphragm should be focussed exactly in the plane of, and on, the thermo-relay strip.

Experimental Method.

1. *Experiments with Galvanometer Mirror Fixed.*—The measurement of the extremely small currents characteristic of the Brownian movement is accompanied by numerous difficulties arising from disturbing influences of all sorts from external and parasitic effects. The elimination of such effects inherent in the relay and the light system itself, as apart from the additional circuits employed, was the first requirement. In short, it is necessary when using the first galvanometer mirror fixed, to obtain a straight line base or registration line for the light spot of the second galvanometer. A preliminary investigation to this end was carried out. The initial registrations obtained were far from linear and contained numerous irregularities.

It was ascertained that the employment of a 12-volt, gas-filled, half-watt lamp with the filament disposed horizontally, gave better results than those obtained for the case of a vertically disposed filament.

The thermo-relay was firmly fixed in a metal case with a fine screw movement in the direction of the thermofoil strip. In order to ensure stable temperature conditions the case was enclosed as far as possible in cotton wool.

Attempts were made to cut down disturbances attributed to air currents, by placing evacuated tubes along the course of the beam of light. Since the tubes introduced no observable improvement they were discarded.

Results taken in the evening when the building was quiet were found to be much better than those obtained during the day, but there still remained disturbances which were attributable to the wind and to seismic movements. It resulted, therefore, that finally the experiments were only carried out after

determining from the neighbouring meteorological station the state of the earth movements.

Additionally, however, a correction was applied for the effect of such movements on the Brownian movement records, for it should be noted that the inevitable slight disymmetry of the galvanometer coil involves a response to such seismic movements.

2. *Experiments with the Galvanometer Mirror Free. Thermocouple effects.*—The above investigations completed, the galvanometer mirror was released and experiments were conducted in the normal way. The first difficulty encountered was that of E.M.Fs. set up in the junctions of the circuit. These effects exist, of course, because of differences of composition of the junctions, which were never precisely the same, so that differences of temperature gave rise both to permanent and to erratic currents. Such effects were minimised by soldering all the leads into brass blocks maintained in thermal contact by employing only mica or thin pieces of parafined paper as the insulating material separating them. The blocks were clamped together in small presses and were covered with cotton wool to protect them from temperature changes, their large heat capacity ensuring uniformity of temperature at the actual junctions.

By employing the above methods a very considerable improvement was effected as was shown by the steadiness of the registration on the rotating drum, and the comparative absence of creep.

Electrostatic effects.—A serious source of error was discovered in the electrostatic effects due to the motion of charges (as of people, especially if on rubber soles) in the neighbourhood of the system and more especially of the outer coils. To avoid such disturbances lead-covered earthed cable was used as leads, and the remaining parts were covered with earthed tinfoil. This arrangement proved perfectly satisfactory, but to make doubly sure all registrations were made with no-one in the room.

To prevent undue conduction of heat into the liquid air, in which one of the coils was immersed, the coil and the part of the leads actually in the liquid air bath was not covered with the tinfoil, the containing vessel being covered instead.

Electromagnetic effects.—The motion of pieces of steel or iron in the vicinity of the galvanometer caused deflections of the light spot by electromagnetic induction. Such movements were of course avoided, and the unavoidable variations of the earths field, etc., were much too small to bring about observable deflections.

Effects peculiar to the liquid air-coil.—The air-coil was well insulated by a

wrapping of cotton wool and introduced no disturbing influences, but the case of the coil immersed in liquid air was different. It was found that owing to the condensation of water vapour at part of the leads, a large and variable electrolytic or battery effect was set up. This, however, was done away with when the coils and leads were carefully insulated with red wax.

The gradient of temperature along the leads from the liquid air gave because of disymmetry, a permanent deflection of the second galvanometer. This was reduced to a small value by carefully heat-insulating the leads with cotton wool and by allowing the coil to be in liquid air for about an hour before observations were made. Thus a state of equilibrium was attained.

All the above described effects were investigated by taking photographic records on the rotating drum, and in this manner it was possible, step by step, to locate the disturbances and to suppress them.

Investigation of the Brownian Movement.

Fig. 2 indicates the circuit used in the measurements. The procedure adopted was as follows:—

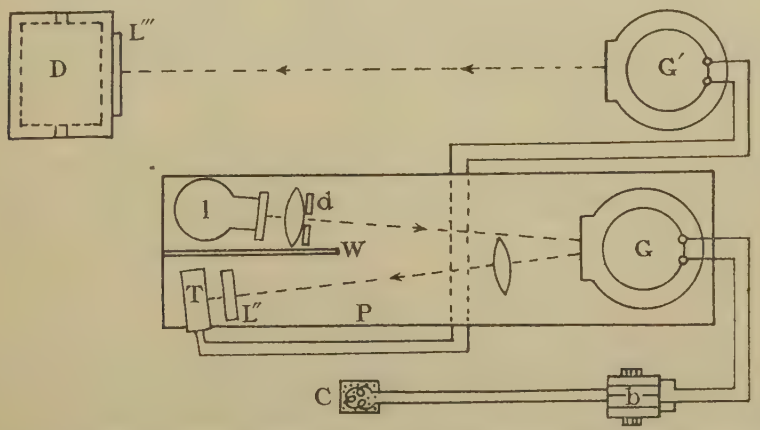


Fig. 2.—General Scheme of Apparatus. *l*, lamp; *d*, diaphragm; G, first galvanometer; L″, cylindrical lens; T, thermo-element in case on metal arm with fine screw for lateral bending; P, heavy metal base plate; G′ second galvanometer; L‴, cylindrical lens condensing line image at drum into a spot; D, recording drum; *b*, brass blocks in clamp; C, coil, in wax, in brass box; W, wadded screen. The first galvanometer circuit is in cotton wool with earthed leadfoil cover.

The sensitivity of the relay was measured before and after each registration. The Brownian movement of the galvanometer alone, on closed circuit, first, and that for open circuit second. The air-coil was then connected to the galvano-

meter by means of brass blocks (fig. 1) which were then clamped-up and brought to room temperature by a current of air. The blocks were then carefully covered up with cotton wool and the system was allowed to rest until no appreciable movement of the spot of light on the registering drum could be detected. The registration was then made.

After this observation the liquid air-coil was connected up in an exactly similar manner, and when conditions attained stability a registration was made, the sensitivity being taken before and after as in the previous registrations. A repeat with the air-coil was then made. The speed of rotation of the drum giving maximum clearness of the Brownian movement record was employed, this being determined experimentally previously.

Details of the Photographic Treatment.

The print of the Brownian movement of the galvanometer coil as obtained on the paper covering the revolving drum may be treated in a variety of ways, but the following method was developed and proved to be most satisfactory.

A slide of the registrations was constructed and projected on squared paper in such a way as to obtain a convenient magnification. Traces of the Brownian movement curves were then drawn out upon the squared paper by hand. In this manner a permanent record was obtained that could be measured-up at leisure and in which the ruled lines could be used for the purpose of base-lines and time intervals.

Equally satisfactory was the method of using an episcope for projecting and magnifying the registration curves, this method of course dispensing with one stage in the previous operations.

Consideration of Results.

The amplitude of the Brownian movement.—In order to find the influence of the cooling of the outer resistance on the amplitude of the Brownian movement it is necessary to calculate from the traced curve a quantity that is the right measure of this amplitude. For the movement of the coil itself we may take the mean value (R.M.S.)

$$\sqrt{\overline{\theta^2}}$$

or the quantity that only differs by a constant factor from this (A.M.)

$$|\overline{\theta}|.$$

It is impossible however to find the real value of θ at any moment, as the lag of the systems thermo-relay-second-galvanometer causes a decrease of the

amplitude of the second galvanometer as compared with the ideal case of an infinitely quick indicator. This difficulty is not as serious as it seems to be,

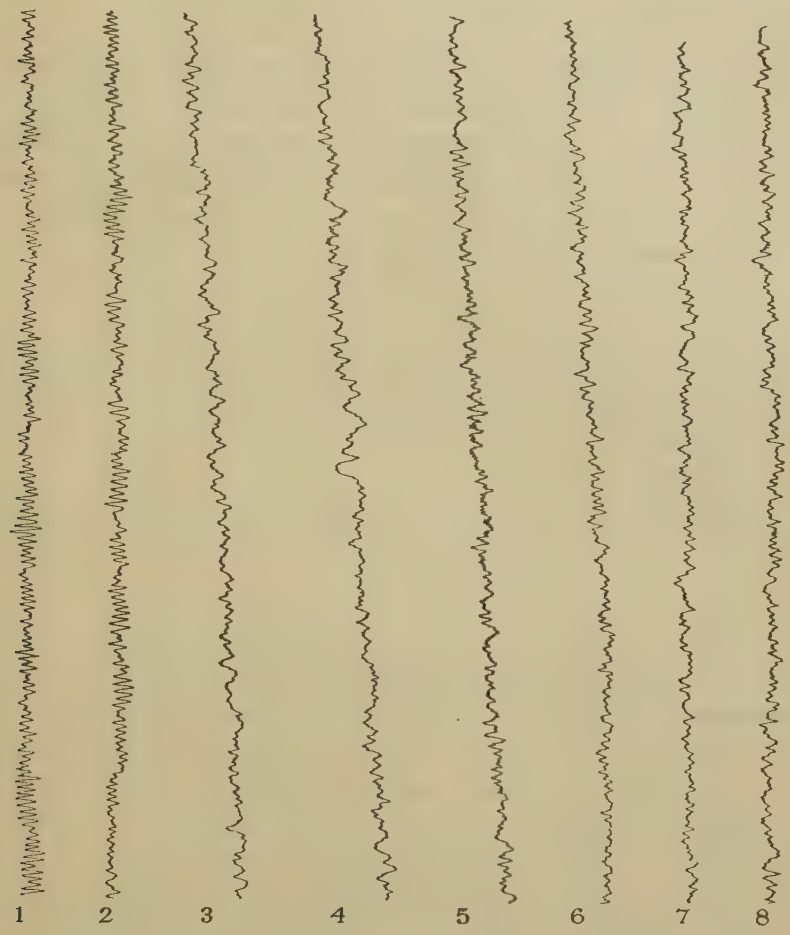


FIG. 3.—Brownian Movement of the Galvanometer Coil. 1 and 2, galvanometer on open circuit; 3 and 4, galvanometer on closed circuit; 5 and 6, coil at air temperature in circuit; 7 and 8, coil at liquid oxygen temperature in circuit. Resistance coils of 130 ohms.

for the amplitude of the second galvanometer is decreased in a ratio that depends only on the mean quickness of variation of θ. This variation is exactly the same in the cases, both of the cooled and of the uncooled outer resistances, and therefore a relative measurement of the amplitude of the Brownian movement can be made from the traced curves.

When we call y the deflection of the second galvanometer as indicated on

an enlarged scale by the traced curve we must try to find one of the quantities

$$\sqrt{\overline{y^2}} \quad \text{or} \quad |\overline{y}|.$$

Here we meet the difficulty that in order to find the deflection y at any point of the curve we must know the zero line, *i.e.*, the line giving the mean position of the light spot of the second galvanometer. When there is a "creep" this position is not constant, and large errors can be introduced into a calculation of y. It is better, therefore, to take the difference between the deflections recorded at a constant time interval, as these can be found without knowing the zero line. The quadratic mean of this difference is

$$\overline{|y_1 - y_2|^2} = 2\overline{y^2} - 2\overline{y_1 y_2}.$$

Both quantities $\overline{y^2}$ and $\overline{y_1 y_2}$ will be proportional to the absolute temperature T or to the effective value of this temperature; $\overline{y^2}$ will be independent of the time interval and $\overline{y_1 y_2}$ will increase from zero for a time interval zero to $\overline{y^2}$ for a very large interval.

To avoid troublesome calculations it is much easier in practice to calculate the value.

$$|\overline{y_1 - y_2}|,$$

this differing from $\overline{|y_1 - y_2|^2}$ only by a constant numerical factor. This mean difference is found in the following way. A large number of ordinates at equal intervals along the curve are read off, for zero line an arbitrary line roughly parallel with the real zero line being taken. The differences between the successive ordinates of this series of readings are added as absolute values and by dividing by the number the value of $|\overline{y_1 - y_2}|$ is found.

When there is a "creep" of the galvanometer the accepted zero line is not parallel with the true one and the value of $|\overline{y_1 - y_2}|$ is inaccurate. This error when calculated for these experiments was in general so small that it could be neglected.

The value of $|\overline{y_1 - y_2}|$ was computed for different time intervals, and it was found to be larger for the longer intervals, as is to be expected. By the theory of v. Smoluckowski* the correlation between the values of the angle θ is given by the relation

$$|\overline{\theta_1 - \theta_2}| = a\,(1 - e^{-t/t_0}) \qquad (6)$$

t_0 being a constant depending on the inertia and the frictional force of the galvanometer coil. Though this relation cannot be applied at once to the

* M. v. Smoluckowski, 'Ann. d. Phys.,' vol. 48, p. 1103 (1915).

deflection θ, as the effect of the lag of the thermo-relay and the second galvanometer is a complicated one, the values of $\overline{|y_1 - y_2|}$ given by experiment show that this quantity depends upon the time in about the same way as we must expect for $\overline{|\theta_1 - \theta_2|}$ according to equation 6. The time t_0 is computed to be 1·5 seconds, a very probable value as the time-period of the first galvanometer is 1·5 seconds. Further, we must expect that the correlation between consecutive deflections is strong for times smaller than the period of the first galvanometer, and that it disappears for an interval that is large in comparison with this period.

The effect of cooling the outer resistance.—In practice, two coils of maganin wire were used, one immersed in liquid air. As the resistance of manganin at the temperature of liquid air is 3 per cent. less than at room temperature the resistance at air temperature of the cooled coil was made 3 per cent. larger than that of the other one, so that under working conditions their resistances were equal.

Several measurements were made with coils of 137 ohms and 400 ohms. For all cases the amplitude of the movement was smaller for the liquid-air coil. On a good and quiet day the ratio of the amplitudes of the cooled and uncooled coils (137 ohms) was 0·77. As the result calculated from equation 5 was 0·70 this was very satisfactory.

That the ratio is larger than the calculated value is to be expected, as every error, assuming that each coil is equally affected, will make the ratio approach unity. If the movement was entirely due to errors, independent of temperature, the ratio would be 1·0.

In some cases the result was not so good, the discrepancy probably being caused by errors occasioned by tremors, seismic movements, &c. In accordance with this explanation, the values of the amplitudes with and without liquid air were both larger than for the best observations under good conditions. A large amplitude indicates a large error, and we must expect, as theory predicts, a ratio of the amplitudes that is much nearer unity.

The errors and the Brownian movement are added as squares, so when $\overline{|y_1 - y_2|} = a$ is the amplitude given by experiment, b the expected amplitude of the Brownian movement, and c the amplitude of the errors, that is, the amplitude which would be present if there was no Brownian movement, then

$$a^2 = b^2 + c^2.$$

Now when we accept that the smallest amplitudes ever measured with a certain coil is the real b, then we can calculate from equation 7 the error c in any other

case, the amplitude a observed with that coil being known. We may assume that the same error exists after the short time during which the one coil is being replaced by the other coil at the different temperature, so by subtracting the same error c as given by equation 7 we may correct these amplitudes and find their corrected ratio. All results corrected in this way gave results in good accord with those predicted by theory. With the 400 ohm coils, for example, the mean of the corrected ratios was 0·61 as compared with a predicted value of 0·59. The error c was always smaller than the amplitude of the Brownian movement, though sometimes it was nearly as large.

It is worth noticing that the largest errors were present on windy days or on days following a period of storm, in which case it is evident that the seismic movements of the Continent, set up by the beating of the sea-waves, will account for a large part of these errors. The slightest asymmetry of the galvanometer coil will give a couple between the transverse movement of the galvanometer to which it is rigidly connected at the bottom, and its rotation.

The fact of the influence of temperature on the movement of the light-spot of the galvanometer, and the accordance of the amount of this influence with the theoretical prediction, make it certain that the movement is caused entirely, or for the larger part, by the Brownian movement. The absolute magnitude of the amplitude, though a less certain criterion, confirms this interpretation.

By a sufficient enlargement we may find the fluctuations in the indications of any galvanometer. When these fluctuations are for the larger part due to Brownian movements we may call the galvanometer good ; when a galvanometer (or system of galvanometers) shows the Brownian movement without enlargement, we may call it sensitive ; when it shows a movement that is largely due to errors, we may call it a bad instrument.

On the Reflexion of Sound from a Porous Surface.

By E. T. Paris, D.Sc.

(Communicated by F. E. Smith, F.R.S.—Received March 31, 1927.)

1. *Introduction.*

When sound-waves impinge on a plane surface forming the boundary between two media of different mechanical properties, part of the incident energy is reflected and part is transmitted as a refracted wave, the sum of the energies of the reflected and refracted waves being equal to the energy of the incident wave. The proportion of the incident energy which is reflected depends not only on the mechanical properties of the two media, but also on the angle of incidence of the primary wave. If the first medium is air and the second a liquid or continuous solid, by far the greater part of the incident sound-energy is reflected. For example, if the second medium is water, 99·9 per cent. of the energy of the primary wave is reflected at normal incidence, while if the angle of incidence exceeds 13 degrees the refracted wave becomes evanescent, and total reflexion occurs.*

If the second medium is not continuous, but of a porous nature, a somewhat different state of affairs ensues. The boundary between the two media is no longer well defined, for the first medium (air) can penetrate into the pores of the second medium and reflexion is then accompanied by a motion of the air in and out of the pores. Examples of porous substances of the kind contemplated are the various " acoustic plasters " produced in recent years for the rapid absorption of sound in auditoria, when the reverberation is found to be in excess of that demanded by good taste. Other examples, of a rather different sort, are hairfelt, carpets and similar materials.

The difference in the behaviour of non-porous and porous surfaces lies in the mechanism by which energy is lost from the incident sound-waves at reflexion. If the surface of the second medium is continuous, energy is transmitted into the refracted wave by vibratory movements of the boundary itself, these movements having the same periodicity as the sound in the incident wave. The rate at which energy is lost from the incident wave per second is the integral over the reflecting surface of the average value of the product of the pressure-variation at the boundary and the velocity of the boundary. If the

* Rayleigh, ' Theory of Sound,' vol. 2, p. 81.

second medium is porous, then, under the influence of the pressures developed by the incident wave, a certain amount of air is moved into and out of the pores in each period. If the lateral dimensions of the pores are small, this in-and-out movement of the air is accompanied (owing to viscosity) by a degradation of sound-energy into heat. The rate at which energy is lost from the incident wave is equal to the integral over the reflecting surface of the average value of the product of the pressures developed at the boundary and the rate (volume per unit time per unit area of surface) at which air passes in and out of the pores. There may, in addition, be loss of energy from the incident wave due to a general vibratory motion of the surface of the porous medium, whereby a refracted wave may be generated, but in the cases to which reference will be made later, such motion is, as a rule, of secondary importance.

When sound is reflected from the surface of a porous material, therefore, the sum of the reflected and refracted energies is no longer (if the pores are small) equal to the incident energy, and the difference, which is the amount degraded into heat by the action of the pores, is said to be dissipated or absorbed. Since, however, in most cases of sound-absorption the refracted sound is of minor interest, it is usual to define the absorbed sound as that which is lost at reflexion. Thus, if E_i is the energy-flux in the incident wave and E_r is the energy-flux in the reflected wave, the coefficient of absorption is defined by

$$\alpha = (E_i - E_r)/E_i. \qquad (1.1)$$

The energy fluxes E_i and E_r are here scalar quantities measured in ergs per square centimetre per second.

The theory of the reflexion of sound by porous bodies received attention at the hands of the late Lord Rayleigh. In a paper* published in 1883 he dealt with the case of sound incident normally on a wall presenting a flat surface and perforated by a great number of similar channels, uniformly distributed, and bounded by surfaces everywhere perpendicular to the face. On the supposition that the channels were of circular cross section, and that the expansions and rarefactions within them took place isothermally, equations were deduced by means of which the ratio of the amplitudes of the incident and reflected waves could be calculated. No experimental verification of these calculations has ever been made, although, with a modern apparatus, it should not be a difficult matter.

* "On Porous Bodies in Relation to Sound," 'Phil. Mag.,' vol. 16, p. 181 (1883): 'Sci. Papers,' vol. 2, p. 220.

In a later paper,* written in 1919 and published posthumously, Rayleigh dealt with the case of sound waves incident at any angle on a porous wall, and it appears that, as in the case of ordinary reflexion, the amplitude of the reflected wave depends on the angle of incidence. It was shown that under certain circumstances the reflected wave may vanish at some particular angle of incidence. There will be occasion later to make more detailed reference to the theoretical results obtained by Rayleigh in this paper.

The formulæ for the amplitude of the reflected wave given by Rayleigh cannot be used to calculate the absorption coefficient of any sound-absorbing substance commonly met with. The theory refers to an ideal substance perforated by a number of similar straight pores of uniform cross section and of equal length. The substances found in practice are of a very different kind. The pores in an acoustic plaster, for example, are of very diverse shape and size. The cross sections are constant neither in area nor shape, nor do the pores run straight, but bend about and lead into one another in a most complicated manner. An examination under a microscope of a section through an acoustic plaster indicates that any attempt to calculate the absorbed sound on the lines of the theory developed by Rayleigh would be of little value. It has been necessary, therefore, to rely entirely on experiment for the determination of the absorption coefficient.

Even when this coefficient has been found experimentally at some particular angle of incidence (and for a given wave length) there has so far been no method by which the absorption could be calculated at any other angle of incidence. Rayleigh's theory has been of no assistance here, for, although it indicates that absorption varies with the angle of incidence, it can only be employed to determine the variation in the special cases to which it refers.

It is the purpose of the present paper to show how an advance may be made by introducing into the theory of the reflexion of sound from a porous wall a quantity called—by analogy with the "acoustical impedance" of A. G. Webster†—the "acoustical admittance per unit area of absorbing surface." This "acoustical admittance" is a quantity which can be readily determined by experiment in a manner to be described, and when it has been found for a particular substance (at a given wave length) the coefficient of absorption for sound of this wave length can be calculated for any angle of

* "On Resonant Reflexion of Sound from a Perforated Wall," 'Phil. Mag.,' vol. 39, p. 225 (1920); 'Sci. Papers,' vol. 6, p. 662.

† 'Proc. Nat. Acad. Sci.,' vol. 5, p. 275 (1919).

incidence by means of simple formulæ. The results obtained appear to be quite general and applicable to any kind of sound-absorbing substance.

2. *Acoustical Admittance.*

The conception of acoustical impedance was introduced by Webster* in 1919, and it has since been developed and applied to the solution of a variety of acoustical problems by G. W. Stewart† and others.‡ According to Webster's original definition the impedance of an acoustical receiver is the (complex) ratio of the excess pressure (due to external sources) at the receiving surface (or orifice) of the receiver to the volume of air periodically displaced by the receiving surface (or moved in and out through the orifice). Stewart, however, has found it more convenient to employ, instead of the ratio of the pressure to the volume-displacement, the ratio of the pressure to the *rate of* volume displacement. Thus, to take an example, suppose that the receiver is a Helmholtz resonator, and let δp be the excess pressure at the mouth of the resonator due to an external source. Also let dq/dt be rate of volume displacement (cubic centimetre per sec.) of air into the resonator through its orifice. Then according to Stewart's definition, the impedance of the resonator is $\delta p/(dq/dt)$. That is $(dq/dt) = \delta p/\mathrm{Z}$, where Z is the impedance. If a velocity-potential, ϕ, is used, $\delta p = \rho\,(d\phi/dt)$, ρ being the density of air, and hence $\mathrm{Z} = \rho\,(d\phi/dt)/(dq/dt) = \rho\phi/q$, ϕ and q being proportional to e^{imt}. In what follows a velocity-potential will always be used, and hence, to avoid an unnecessary repetition of ρ, the above definition will be modified and impedance will be defined as the ratio ϕ/q, this being equal to Stewart's impedance divided by ρ. It is also equal to $(\rho d\phi/dt)/(\rho dq/dt)$, or pressure to mass-flow. Its dimensions are $\mathrm{L^{-1}\,T^{-1}}$, compared with $\mathrm{ML^{-4}\ T^{-2}}$ for Webster's, and $\mathrm{ML^{-4}\ T^{-1}}$ for Stewart's impedance.

It will, moreover, be convenient to employ, instead of impedance as thus defined, its reciprocal, which will be denoted by Ω. Thus $q = \Omega\phi$. For con-

* *Loc. cit.* The paper was actually read before the American Physical Society in Philadelphia in 1914, but publication was delayed.

† Stewart's work is contained in a series of papers in the 'Physical Review.' The more important are: vol. 16, p. 313 (1919) on conical horns; vol. 20, p. 528 (1922), on acoustic wave-filters; vol. 26, p. 688 (1925), vol. 27, pp. 487 and 494 (1926), on the transmission of sound through pipes with branch-lines; and vol. 28, p. 1038 (1926), on the measurement of acoustical impedance.

‡ Notably Kennelly in connection with the theory of the telephones. See his 'Electrical Vibration Instruments,' chap. xiii (1923).

venience, Ω may be called the "acoustical admittance."* On Stewart's definition, admittance would, of course, be equal to Ω/ρ.

In order to make use of this conception in the theory of the reflexion of sound from the surface of a porous substance consider the case of plane waves incident on a porous wall of infinite extent. The velocity-potential at the surface of the wall will be the sum of the potentials of the incident and reflected waves, say $(\phi + \phi')$. In an element of surface σ, of dimensions small compared with the wave-length, a certain volume (Q) of air will move periodically in and out of the pores of the wall. Thus, by analogy with the case of the acoustical receiver considered above, we may write $Q = \Omega\,(\phi + \phi')$, where Ω is the admittance of the area σ. If q is the volume displaced per unit area of the wall, $q = \Omega'\,(\phi + \phi')$, where Ω' is written for (Ω/σ), and may be called the *admittance per unit area* of the reflecting surface. The dimensions of Ω' are those of the reciprocal of a velocity ($L^{-1}T$).

Physically, the equation $q = \Omega'\,(\phi + \phi')$ implies a linear relation between the amplitude of the velocity-potential (or the pressure) at the surface of the wall and the volume-displacement (or the mass-flow) of air into the pores. In the case of the small amplitudes with which we are concerned, such an assumption appears to be quite justifiable. Also, it will be assumed that Ω' is independent of θ, which implies that disturbances are not propagated laterally, from one part of the porous medium to another, to any appreciable extent.

3. *Reflexion from a Porous Wall.*

Let the reflecting surface of the wall lie in the plane $x = 0$, and let plane waves of sound with wave-fronts parallel to the axis of z be incident upon it at any angle θ. The velocity-potential of the incident waves supposed travelling towards $-x$, may be represented by

$$\phi = Ae^{ik(at + x\cos\theta + y\sin\theta)}, \tag{3.1}$$

and that of the reflected waves by

$$\phi' = Be^{ik(at - x\cos\theta + y\sin\theta)}, \tag{3.2}$$

where $k = 2\pi/\text{wave-length}$ and a is the velocity of sound in free air. At $x = 0$, ϕ and ϕ' must satisfy the equation

$$\frac{\partial}{\partial x}(\phi + \phi') = \Omega'\frac{\partial}{\partial t}(\phi + \phi'), \tag{3.3}$$

* In electrical science admittance is used to denote the reciprocal of reactance, that is, the imaginary part only of a complex impedance.

since each side of this equation is equal to the current per unit area (dq/dt) into the pores of the wall. By substituting in (3.3) the values of ϕ and ϕ' given in (3.1) and (3.2), it is found that

$$A - B = \frac{a\Omega'}{\cos\theta}(A + B), \tag{3.4}$$

or

$$\frac{B}{A} = \mu = \frac{\cos\theta - a\Omega'}{\cos\theta + a\Omega'}. \tag{3.5}$$

The ratio B/A, or μ, may be called the coefficient of reflexion.

The energy-flux in the incident wave is proportional to $|A|^2$ and that in the reflected wave to $|B|^2$. Hence the coefficient of absorption is

$$\left.\begin{aligned} \alpha = \frac{E_i - E_r}{E_i} &= \frac{|A|^2 - |B|^2}{A|^2} \\ &= 1 - |\mu|^2 = 1 - \left|\frac{\cos\theta - a\Omega'}{\cos\theta + a\Omega'}\right|^2 \end{aligned}\right\} \tag{3.6}$$

It is evident from (3.5) and (3.6) that when Ω' has been determined the coefficients of reflexion and absorption can easily be calculated for any value of the angle of incidence.

It is of interest to compare these results with those given by Rayleigh in his later paper.* Rayleigh considered the case of a perforated wall in which the perforated and unperforated areas were in the proportion σ/σ'. The channels were supposed to be uniformly distributed and of length l and dissipation of acoustical energy was represented by the introduction of a term $h(d\psi/dt)$ in the equation to be satisfied by the velocity-potential ψ within the channels. Rayleigh's equation (15) shows that under these conditions :—

$$\frac{B - A}{B + A} = \frac{\sigma}{(\sigma + \sigma')\cos\theta}\frac{k' \tan k'l}{\imath k}, \tag{3.7}$$

where $k = 2\pi/\lambda$, as before, and $k'^2 = k^2 - \imath nh/a^2$, n being 2π times the frequency of the incident sound. From equation (3.4) we have

$$\frac{B - A}{B + A} = -\frac{a\Omega'}{\cos\theta}. \tag{3.8}$$

Thus it will be seen that cos θ appears in Rayleigh's equation (15) in just the way that would be expected from the theory of reflexion from a porous wall as developed with the aid of acoustical admittance. Also (3.8) becomes identical with (3.7) if

$$\Omega' = -\frac{1}{a}\frac{\sigma}{\sigma + \sigma'}\frac{k' \tan k'l}{\imath k}. \tag{3.9}$$

* *Loc. cit.* (1920).

This, therefore, is the theoretical value of the acoustical admittance per unit area appropriate to the problem considered by Rayleigh.

Two special cases noted by Rayleigh are instructive. The first is when there is no dissipation (or absorption), so that, with $h = 0$, $k'^2 = k^2$. The amplitude (A) of the incident wave being supposed equal to unity, Rayleigh's equation (17) gives for the reflected wave

$$\mathrm{B} = \frac{\cos\theta\cos kl - \iota\sin kl}{\cos\theta\cos kl + \iota\sin kl}, \tag{3.91}$$

so that mod B = 1. Comparing with (3.5) we see that

$$a\Omega' = \iota\tan kl. \tag{3.92}$$

So that, in this case, Ω' is entirely imaginary. The second case is when the channels are so long that the vibrations within them are sensibly extinguished before the stopped end is reached. The incident amplitude being again supposed unity, Rayleigh's equation (20) for the reflected amplitude is

$$\frac{\mathrm{B}-1}{\mathrm{B}+1} = -\frac{\sigma}{(\sigma+\sigma')\cos\theta}, \tag{3.93}$$

whence, by (3.8)

$$a\Omega' = \sigma/(\sigma+\sigma'), \tag{3.94}$$

Ω', and therefore μ, is in this case real, and there is no change of phase at reflexion.*

In the second case there is always some value of θ for which $\cos\theta = \sigma/(\sigma+\sigma')$, and when this is the case the reflected wave vanishes. Thus, if the perforated and unperforated areas are equal, $\sigma/(\sigma+\sigma') = \frac{1}{2}$, and the reflected wave vanishes when $\theta = 60^\circ$. At this angle, therefore, the incident sound is completely absorbed.

4. *Method of Determining* Ω'.

To determine Ω' for any particular substance it is necessary to find the ratio of the amplitudes of the incident and reflected waves at some given angle of incidence and the change of phase at reflexion. If the reflecting face of the substance lies in the plane $x = 0$, and plane waves, travelling towards $-x$, impinge on it normally, the velocity-potential of the incident and reflected waves may be represented by

$$\left.\begin{aligned}\phi &= \mathrm{A}\varepsilon^{\iota kx}\,.\,\varepsilon^{\iota kat}\\ \phi' &= \mathrm{B}\varepsilon^{-\iota k(x+\delta)}\,.\,\varepsilon^{\iota kat}\end{aligned}\right\} \tag{4.1}$$

where A and B may be taken to be real since the factor $\varepsilon^{-\iota k\delta}$ has been introduced

* By "change of phase at reflexion" is meant the angle $\tan^{-1}(\mu_2/\mu_1)$ where μ_2 and μ_1 are the imaginary and real parts of μ respectively.

to allow for the change of phase at reflexion. If Ω' is the acoustical admittance per unit area of the substance, we have, corresponding with (3.3)—

$$\frac{\partial}{\partial x}(\phi + \phi') = \Omega' \frac{\partial}{\partial t}(\phi + \phi') \text{ when } x = 0. \tag{4.2}$$

By substitution from (4.1) we find—

$$A - B\varepsilon^{-\iota k\delta} = a\Omega'(A + B\varepsilon^{-\iota k\delta}), \tag{4.31}$$

or if $\mu_0 = (B/A)$,*

$$1 - a\Omega' = \mu_0\varepsilon^{-\iota k\delta}(1 + a\Omega'). \tag{4.32}$$

Let $\Omega' = \Omega'_1 + \iota\Omega'_2$. Then, equating the real and imaginary parts of (4.32), we obtain—

$$\left.\begin{aligned} a\Omega'_1 \,.\, (1 + \mu_0 \cos k\delta) + a\Omega'_2 \,.\, \mu_0 \sin k\delta - (1 - \mu_0 \cos k\delta) = 0 \\ a\Omega'_1 \,.\, \mu_0 \sin k\delta - a\Omega'_2 \,.\, (1 + \mu_0 \cos k\delta) + \mu_0 \sin k\delta = 0 \end{aligned}\right\} \tag{4.4}$$

Whence

$$\left.\begin{aligned} a\Omega'_1 &= \frac{1 - \mu_0^2}{1 + 2\mu_0 \cos k\delta + \mu_0^2} \\ a\Omega'_2 &= \frac{2\mu_0 \sin k\delta}{1 + 2\mu_0 \cos k\delta + \mu_0^2} \end{aligned}\right\} \tag{4.5}$$

If, therefore, μ_0 and $k\delta$ can be found experimentally Ω' can easily be calculated with the help of (4.5), the velocity of sound (a) being assumed to be known.

A convenient means of finding μ_0 and $k\delta$ is provided by the "stationary wave" method of measuring sound-absorption at normal incidence. In this method, due originally to Tuma (1902),† the specimen of sound-absorbent substance to be tested is provided in the form of a flat disc and is used to close one end of a cylindrical pipe, the other end of which remains open (see fig. 1). Opposite to the open end is placed a source of sound, waves from which pass down the pipe and are reflected back from the specimen. When a steady state is reached, the motion of the air-particles within the pipe is the same as that due to a single train of incident and a single train of reflected waves. It is easy to see that under these conditions there will exist in the pipe a series of positions of maximum and minimum pressure-amplitude, the maxima alternating with the minima at distances of a quarter wave length. Also, if A and B are the amplitudes of the potentials of the incident and reflected waves, the maximum pressure-

* μ_0 is the modulus of the coefficient of reflexion at normal incidence.

† 'Sitz. K. Akad. Wiss. Wien,' 2A, vol. 3, p. 402 (1902). The method, in modified forms, has been used by Weisbach, 'Ann d. Physik,' vol. 33, p. 763 (1910); Taylor, 'Phys. Rev.,' vol. 2, p. 270 (1913), and Eckhardt and Chrisler, 'Sci. Paper of the Bureau of Standards,' No. 526 (1926).

amplitude will be proportional to (A + B) and the minimum to (A − B). It is the ratio of (A + B) to (A − B) which can be measured experimentally (in the manner indicated in the next section) and from this observed ratio the value of Ω' is readily deduced. Thus, if

$$e/f = (\mathrm{A} + \mathrm{B})/(\mathrm{A} - \mathrm{B}), \tag{4.61}$$

we have

$$\mu_0 = \frac{(e/f) - 1}{(e/f) + 1}. \tag{4.62}$$

The velocity-potentials of the incident and reflected waves within the pipe can be represented as in (4.1), the total potential being the sum of ϕ and ϕ'. Thus, retaining only the real parts, we have

$$\phi + \phi' = \mathrm{A} \cos k\,(at + x) + \mathrm{B} \cos k\,(at - x - \delta). \tag{4.7}$$

Let

$$t' = t - \tfrac{1}{2}\,(\delta/a), \quad x' = x + \tfrac{1}{2}\delta.$$

Then

$$\begin{aligned}\phi + \phi' &= \mathrm{A} \cos k\,(at' + x') + \mathrm{B} \cos k\,(at' - x') \\ &= (\mathrm{A} + \mathrm{B}) \cos kx' \cos kat + (\mathrm{A} - \mathrm{B}) \sin kx' \sin kat'. \end{aligned} \tag{4.8}$$

So that the motion within the pipe may be looked upon as being due to two superimposed stationary waves, one of amplitude (A + B) and the other of amplitude (A − B), the nodes of the former being at the loops of the latter and *vice versa*. The maximum and minimum values of $\phi + \phi'$ (and therefore of the pressure-amplitude), that is (A + B) and (A − B), occur when $\cos kx' = \pm 1$ and $\cos kx' = 0$, respectively. If the sound-absorbing specimen be removed and a rigid reflecting plate substituted, there will be complete reflexion without change of phase, so that A = B and $\delta = 0$. The velocity-potential within the pipe is then

$$\phi + \phi' = 2\mathrm{B} \cos kx \cos kat, \tag{4.9}$$

this being the potential of a stationary wave with nodes and loops occurring when $\cos kx = \pm 1$ and $\cos kx = 0$, respectively. We see, therefore, by comparing (4.8) with (4.9), that the positions of the maxima and minima (when the absorbing specimen is in place) differ from the positions of the nodes and loops (when a reflecting plate is used) by a distance $x' - x = \frac{1}{2}\delta$. By means of the apparatus described in the next section the positions of the maxima and minima and nodes and loops can be determined with considerable accuracy and in this way δ is measured directly. When the reflecting plate is in position, the distance from loop to loop (*i.e.*, $\frac{1}{2}\lambda$) can also be measured accurately, so that $k\delta$ $(= 2\pi\delta/\lambda)$ is fully determined.

5. *Experimental Arrangements.*

The form of stationary wave apparatus employed for the measurement of Ω' is described elsewhere* in some detail in connexion with the measurement of sound-absorption at normal incidence, and it is only necessary here to refer briefly to the principal features. The apparatus, which is shown diagrammatically in fig. 1, was designed for use at frequencies between 200 and 600 vibrations per second. The dimensions of the pipe are 30·5 cm. diameter and 226

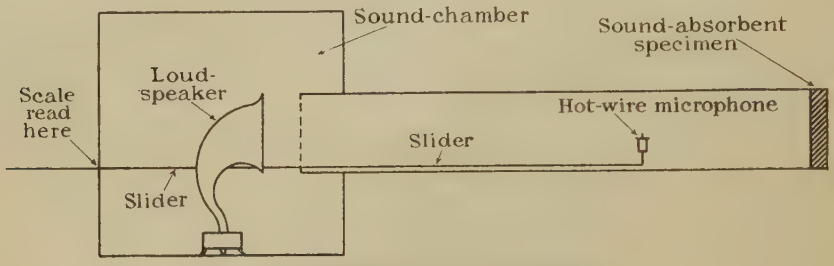


Fig. 1.—Stationary Wave Apparatus.

cm. length. The source of sound is a loud-speaker driven by a thermionic valve oscillator, the loud-speaker being fitted with a special device by which the strength of the source can be maintained constantly at the same value. The open end of the pipe and the source of sound are enclosed in a sound-chamber to prevent the strength of the source being influenced by the movements of the observer. For exploring the sound inside the pipe, use is made of a small form of hot-wire microphone. The microphone is carried at the end of a slider which lies along the bottom of the pipe, and passes out through a slot in the wall of the sound chamber. The distance of the microphone from the closed end of the pipe is found by means of a millimetre scale on the slider, the scale being read where the slider passes through the wall of the sound-chamber.

The response of the hot-wire microphone is measured by the ohmic change in its electrical resistance. To determine the ratio of the maximum to the minimum pressure-amplitude (*i.e.* $(A + B)/(A - B)$), when the absorbent specimen is in place, the following method, which eliminates the relation between the response of the microphone and the amplitude of the sound affecting it, is used. With the specimen in place, the minimum and maximum ohmic changes, say ρ_1 and ρ_2, are found by moving the microphone along the pipe. The specimen is then removed and replaced by a disc of brass which causes almost complete reflexion without change of phase. The microphone is first of all placed at a loop position in the resultant stationary wave (where there is no change in its

* In a paper submitted to the Physical Society of London (1927).

electrical resistance), and then moved along the pipe until the resistance changes ρ_1 and ρ_2 are observed. The distances from the loop at which these changes occur, say y_1 and y_2, are noted. Since the pressure-amplitude in the stationary wave is proportional to sin ky, when y is measured from a loop, it follows that the ratio of maximum to minimum pressure-amplitude is sin ky_2/sin ky_1. Thus

$$\mu_0 = \frac{(e/f) - 1}{(e/f) + 1} = \frac{\sin ky_2 - \sin ky_1}{\sin ky_2 + \sin ky_1}.* \qquad (5.1)$$

To determine $k\delta$ it is only necessary to observe the distances of, say, the first minimum from the absorbent specimen and the first loop from the brass reflector, the difference between these two distances being equal to $\frac{1}{2}\delta$. The distance of the first loop to the second loop gives the half wave-length of the sound, and hence $k\delta$ is completely determined.

To ascertain the position of a minimum or loop accurately, it is best to find two positions, one on each side of the minimum or loop, at which the electrical resistance of the microphone shows equal small changes. The mean of the scale-readings at these two positions gives the scale-readings for the loop or minimum. With the apparatus described above the position of a loop or minimum can be determined in this way without difficulty to within 1 mm. when the wave-length is 66 cm.

6. *The Variation of Coefficient of Absorption with Angle of Incidence in the case of an Acoustic Plaster.*

With the aid of the apparatus described in the preceding section, observations were made on a specimen of an acoustic plaster prepared at the Building Research Station of the Department of Scientific and Industrial Research. The method of manufacturing this plaster is the subject of a patent held by the Building Research Board. The constituents are magnesium chloride, granulated slag, magnesium oxide, lime, powdered aluminium and glue. These are beaten up to a froth, the formation of which is assisted by the evolution of gaseous bubbles by the reaction between the lime and the powdered aluminium. The included bubbles give the plaster its porous and sound-absorbent character.

With sound of wave-length 66 cm. (about 512 vibrations per second) it was found that the ratio (A + B)/(A − B), or e/f, was 12·4. Hence :—

$$\mu_0 = \frac{(e/f) - 1}{(e/f) + 1} = \frac{11{\cdot}4}{13{\cdot}4} = 0{\cdot}851.$$

* A small correction is needed to allow for the effect of the presence of the microphone on the motion of the air within the pipe.

Also that $\frac{1}{2}\delta$ was $-2\cdot2$ cm., the negative sign denoting that the distance of the minimum from the absorbing specimen was $2\cdot2$ cm. less than the distance of the corresponding loop from the reflecting plate, that is, $2\cdot2$ cm. less than a quarter wave-length. Hence.

$$k\delta = -\frac{180 \times 2 \times 2\cdot2}{66} = -11\cdot5^\circ,$$

so that $\cos k\delta = 0\cdot980$, $\sin k\delta = -0\cdot199$.

By means of the formulæ (4.5) $a\Omega'$ can now be calculated. It is found that

$$a\Omega' = 0\cdot0813 - \iota \times 0\cdot0100.$$

We are now in a position to calculate the coefficient of absorption (α) for any given angle of incidence. We have

$$\mu = \frac{\cos\theta - a\Omega'}{\cos\theta + a\Omega'},$$

and since it is the square of the modulus of μ in which we are interested (since $\alpha = 1 - |\mu|^2$) we may calculate directly

$$1 - |\mu|^2 = 1 - \frac{(\cos\theta - a\Omega_1')^2 + (a\Omega_2')^2}{(\cos\theta + a\Omega_1')^2 + (a\Omega_2')^2}, \tag{6.1}$$

where $a\Omega_1' = 0\cdot0813$ and $a\Omega_2' = -0\cdot0100$, so that

$$\alpha = 1 - |\mu|^2 = 1 - \frac{(\cos\theta - 0\cdot0813)^2 + 0\cdot0100}{(\cos\theta + 0\cdot0813)^2 + 0\cdot0100}.$$

The results of the calculation are remarkable, and are shown graphically by the curve in fig. 2. At normal incidence ($\theta = 0^\circ$) the coefficient of absorption

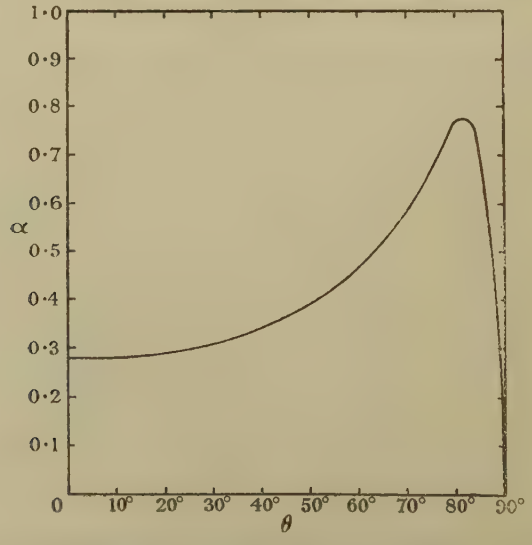


Fig. 2.—Variation of Coefficient of Absorption (α) with Angle of Incidence (θ) for an Acoustic Plaster. $= 66$ cm.

is 0·28. As θ is increased the absorption increases, slowly at first, and then more rapidly, passing through a maximum value of 0·76 when $\theta = 83°$. Thereafter the absorption drops rapidly to zero at grazing incidence. At no angle of incidence between 0° and 89° is the absorption less than it is at normal incidence.

The results obtained with this specimen of plaster are typical of the behaviour of materials of this kind, of which a considerable number have been examined. The change of phase ($k\delta$) is in all cases found to be small, and hence, the imaginary part of Ω', as shown by (4.5), is also small. It follows, from (6.1), that there is always some angle of incidence, approximately given by $\cos\theta = a\Omega_1'$ at which there is heavy absorption.

If a material were found for which $k\delta$ was zero, there would necessarily be some angle of incidence at which there would be complete absorption. This is, of course, the case with the ideal substance considered by Rayleigh, for which $a\Omega' = \sigma/(\sigma + \sigma')$.

A direct experimental verification of these results would be of great interest. The most promising method would appear to be that used by Watson,* in which a beam of sound is produced by a source placed at the focus of a paraboloid. The beam is directed on to the specimen and the amplitudes of the incident and reflected sounds compared by direct measurement with a Rayleigh disc instrument. The angle of incidence could be varied by rotating the specimen On account of the wave length of sound the apparatus required would be somewhat large.

The investigation described above formed part of a research into the methods of measuring sound-absorption, undertaken in the Acoustical Section of the Air Defence Experimental Establishment, with the aid of funds provided by the Department of Scientific and Industrial Research, on the recommendation of their Physics Research Board.

* F. R. Watson, 'Acoustics of Buildings,' chap. 8 (1923).

The Efficiency of K Series Emission by K Ionised Atoms.

By L. H. Martin, M.Sc., 1851 Exhibition Scholar (Melb.).

(Communicated by Sir Ernest Rutherford, P.R.S.—Received April 28, 1927.)

§ 1. *Introduction.*

The general outlines of the atomic processes resulting from an initial state of ionisation, with the subsequent emission of a quantum of radiation are well known.

In the particular case of the excitation of K series radiation by X-rays of frequency ν, the initial ionisation of the atom is associated with the ejection of an electron with a velocity given by the familiar Einstein-Bohr relation

$$\tfrac{1}{2}mv^2 = h[\nu - \nu_K]. \qquad (1)$$

In the reorganisation of the atom an electron from an outer level takes up the orbit in the K shell vacated by the K electron, the resulting excess energy of the atom appearing as radiation of frequency $\nu_{K\alpha\beta\gamma}$ given by the relations

$$\nu_{K\alpha\beta\gamma} = \nu_K - \nu_{LMN,\,etc.}, \qquad (2)$$

where ν_K, $\nu_{LMN,\,etc.}$, are quantities corresponding to the energies $h\nu_K$, $h\nu_{LMN,\,etc.}$, of the atom in its initial and final states.

The essential adequacy of this picture seems now to be well established. In particular this is evidenced by the striking success with which Kossel, Sommerfeld, Wentzel, Bohr, Coster and Dauvillier* have been able to correlate the frequencies of the constituent radiations of the various series in X-ray spectra.

However, it has been apparent for some time that this simple model does not give a complete account of the phenomena associated with the reorganisation.

In some early measurements on the efficiency of excitation by X-rays of K characteristic radiation in atoms belonging to the Fe group, Barkla and Sadler† found that only a small fraction, some 30 per cent., of these atoms which were K ionised emitted K series radiation. Kossel‡ has suggested that this low efficiency of K series emission may be explained on the assumption that the majority of such atoms undergo radiationless reorganisations of the Klein

* ' Atombau und Spektrallinien,' 4th Ed., p. 303.

† Barkla, Bakerian Lecture, ' Phil. Trans.,' vol. 217, p. 315 (1917).

‡ ' Z. f. Physik,' vol. 19, p. 333 (1923).

Rosseland* type resulting in the appearance of high speed electrons with energies E given by the equation—

$$E = h\,[\nu_{K\alpha\beta\gamma} - \nu_{LMN,\,etc.}]. \qquad (3)$$

Recently P. Auger,† from a study of the tracks produced in a Wilson expansion chamber containing different inert gases which were ionised by X-rays of known wave-length, has been able to demonstrate definitely the existence of such radiationless processes. Auger has photographed examples of "paired" tracks which have a common origin and ranges which correspond to photo-electrons possessing energies given by equations (1) and (3) respectively. It is interesting to note that he has obtained cases where as many as four tracks have a common origin indicating that this radiationless process may continue throughout the outer levels resulting finally in an atom ionised to a high degree in the surface levels.

Bothe‡ has recently suggested an alternative explanation of these results. He points out that the results of Barkla and Sadler can be correlated with those of Auger on the assumption that they are manifestations of the phenomena of "internal absorption" in the province of X-rays, a conception already familar in the picture given by Ellis and Skinner§ of the origin of β ray spectra.

According to this hypothesis K radiation is always emitted in the reorganisation of a K ionised atom, but owing to reasons at present undetermined the probability that this K radiation shall be absorbed in the L, M, N, etc., levels of the same atom is very much greater than that it shall be absorbed in the same levels of some other atom. In other words, the absorption coefficient for the K series radiation in the L, M, N, etc., levels of the "parent" atom is very much greater than that for radiation of the same wave-length emitted from some source outside this atom, with the result that less K radiation is found to be emitted from a metal plate, when this is illuminated with exciting X-rays, as in Barkla's experiments, than the amount deduced using normal values for the absorption coefficients of the fluorescent radiation.

The excitation of the K series radiations in a plate or gas will be accompanied by a sudden increase in the photoelectron emission as a result of the "internal absorption" of the fluorescent radiations in the outer levels of the atoms. The energies of these electrons are again given by equation (3). Experimental evidence of this increase in the emission of high speed electrons

* Klein and Rosseland, 'Z. f. Physik,' vol. 4, p. 46 (1921).

† 'C.R.,' vol. 180, p. 65 (1925); 'J. Phys. Rad.' (June, 1925).

‡ 'Phys. Z.,' vol. 26, p. 410 (1925).

§ 'Roy. Soc. Proc.,' vol. 105, p. 185 (1924).

is strikingly shown by some early measurements by Barkla* and Beatty† of the ionisation produced relative to air in bromine, iodine, and selenium hydride respectively by a series of homogeneous X radiations. They observed a sudden large increase (2 or 3 times) in the ionisation produced when the wave-length of the radiation absorbed was slightly shorter than that of the K limit of the gas in question.‡ Such an increase would not be expected as the dimensions of the ionisation chambers were such that only a negligible amount of excited K radiation was absorbed in the gas and according to equation (1) the photo-electrons ejected from the K level by radiation whose frequency approximates to that of the K limit have zero velocity when they leave the atom.

This unexpected and hitherto unexplained increase in ionisation at the K limit first finds a satisfactory explanation in terms of the " internal absorption " hypothesis.

It is shown later that values of the efficiency of emission of K series radiation for these gases can be deduced from these experiments which agree satisfactorily with other more direct determinations.

It must be mentioned, however, that at present there exists no experimental evidence confirming the assumption that in these radiationless reorganisations the K radiation is actually responsible for the high speed photo-electron (equation 3) which is the only evidence of its presence, as this high speed electron may of course equally well be due to a direct interaction between the K shell and one of the outer shells.

The conception of " internal absorption " provided it is not interpreted too strictly has a real advantage, however, as it presents a possible simple model of the atomic process, and for that reason it will be used throughout this paper as a descriptive term for the radiationless process in the reorganisation of an ionised atom.

It might be expected that atomic processes subsequent to the initial ionisation of the atom in the K shell would be independent of the mechanism responsible for this ionised state ; in particular that the probability of emission of K series radiation as opposed to the probability of its " internal absorption " would be independent of the frequency of the ionising radiation. In Table I some values are shown of " the number of quanta of K characteristic radiation emitted by each K ionised atom " for a number of different atoms excited by a series of

* Barkla and Philpot, ' Phil. Mag.,' vol. 25, p. 849 (1923).

† ' Roy. Soc. Proc.,' A, vol. 85, p. 329 (1911).

‡ See fig. 3, at end of paper.

X radiations of different frequency. These values are deduced from the results of Barkla and Sadler.

Table I.

Exciting Radiations K series.	Nos. of quanta of K radiation emitted per K ionised atom.					
	Cr.	Fe.	Co.	Ni.	Cu.	Zn.
Fe	0·20	—	—	—	—	—
Co	0·20	0·29	—	—	—	—
Ni	0·19	0·28	0·35	—	—	—
Cu	0·18	0·28	0·35	0·39	—	—
Zn	0·15	0·25	0·30	0·33	0·38	—
As	0·13	0·21	0·24	0·29	0·32	0·40
Se	—	0·21	—	—	0·32	0·36
Ag	—	—	—	—	0·33	0·34

It will be seen that the efficiency of K emission decreases steadily as the frequency of the exciting radiation increases. This result is surprising and seemed to be of sufficient importance to merit further experiments under the improved conditions which result from the more detailed knowledge available to-day of X-ray phenomena.

In the present experiment relative values for the efficiency of K emission have been determined in the case of Fe, Ni, Cu and Zn for a series of exciting radiations which lie between 0·6 Å.U. and the respective K limits of these elements. In the particular case of Fe the absolute efficiency has been determined and investigated as a function of the frequency of the ionising radiation.

The experimental details are given in the next section; it is only necessary to say here that a beam of exciting radiation is totally absorbed in a plate of one of the above metals placed at right angles to it. The ratio of the energy E_K of the total excited K radiations to that of the exciting radiation E_E can be deduced readily from the ratio of the ionisations produced by the total absorption in oxygen of the exciting beam and those excited rays which escape the radiator, together with a knowledge of the absorption coefficients for those radiations in the radiator.

Expressing the energies of the excited and exciting radiations as $n_K h\nu_K$, $n_\nu h\nu_E$, respectively, the number of K quanta emitted per ionised atom is given simply by

$$f = \frac{n_K}{n_\nu} = \frac{E_K}{E_E} \cdot \frac{\nu_E}{\nu_K}. \qquad (4)$$

If J_K, the K jump, is the value of the ratio

$$\mu/\rho\ (K + L + M, \text{ etc.})/\mu/\rho\ (L + M + \text{etc.})$$

corresponding to ν_E, the ratio of the number of exciting quanta absorbed in the K level to the number absorbed in the whole atom is $[J_K - 1]/J_K$, and the efficiency of emission ϕ for K ionised atoms is given by

$$\phi = f\,\frac{J_K}{J_K - 1}. \tag{5}$$

As the variation of J_K with ν is of importance, and as the only available values for the absorption coefficients of the different radiators for wave-lengths in the neighbourhood of the K absorption limits are those of Barkla and Sadler* made for radiations which were not strictly homogeneous, some new measurements were made the results of which are given in another paper.†

Recently Balderston* has published an account of measurements of ϕ for these elements and Mo Ag and Sn in a wave length range 0·306 Å.U. to 0·889 Å.U. radiations which are in general considerably shorter than those used here. A comparison of these results with those of the present paper is of particular interest in throwing further light on the variation of ϕ with the frequency of the exciting radiation. At the end of this paper, the values of ϕ obtained for different elements in other experiments are compared in relation to the general variation with atomic number.

§ 2. *Theory of Experiment.*

It has already been mentioned that the value of the energy of the excited K radiation obtained experimentally has to be corrected for the absorption in the metal plate which this radiation undergoes before it escapes the surface. This correction has to be extended to correct for the fact that less fluorescent radiation is excited in the deeper layers than in the surface layers owing to the progressive decrease in intensity of the exciting beam.

If there are n_ν quanta of radiation of frequency ν_E incident on the radiator per second, the number absorbed in a layer δx at a depth x is given by $n_\nu e^{-\alpha_1 x}\,\alpha_1\,\delta x$ where α_1 is the absorption coefficient of this radiation in the radiator. By definition of f the number of K characteristic quanta excited in this layer is $f\,.\,n_\nu\,.\,e^{-\alpha_1 x}\,.\,\alpha_1\,.\,\delta x$; while the number which escape from the surface of the plate in a direction θ is

$$(2\pi \sin\theta\,\delta\theta/4\pi)\,.\,f\,.\,n_\nu\,.\,e^{-\alpha_1 x}\,.\,e^{-\alpha_2 x \sec\theta}\,\delta x,$$

* 'Phil. Mag.,' vol. 17, p. 739 (1909).
† Martin, 'Proc. Camb. Phil. Soc.' (1927).
‡ 'Phys. Rev.,' vol. 27, p. 696 (1926).

where α_2 is the " mean " absorption coefficient of the K characteristic radiation. In these equations it has been assumed that the " true " absorption coefficient can be replaced by the " total " absorption coefficient. This is justified, since the minimum value of μ/ρ which enters into this equation is 27·0 (the maximum error is less than 1 per cent.). Finally, the ratio of the excited to the exciting quanta on integrating is given by

$$\frac{N_K}{n_\nu} = \frac{f}{2}\left[(1-\cos\Omega) - \frac{\alpha_2}{\alpha_1}\log_e\frac{\alpha_1+\alpha_2}{\alpha_1\cos\Omega+\alpha_2}\right]. \qquad (6)$$

In this integration θ has been assumed constant in the x integration. This is justified as the average maximum depth from which the excited rays can escape is of the order of 0·001 cms. Ω the maximum value assumed by θ is an apparatus constant.

It is found that even with large values of Ω ($\Omega \doteqdot 90°$) the ratio of the ionisation currents produced by the absorption of the excited and exciting radiations near the K limit of the radiator in the same length of oxygen is very small (of the order 5 to 10 per cent.). Large values of Ω necessitated the use of a short ionisation chamber in which the electrodes are perpendicular to the X-ray beam. Since, with such a chamber only a small fraction of the X-ray beams are absorbed corrections have to be made for the fact that the excited and exciting radiations possess different absorption coefficients.

In the case of hemispherical electrodes of radii x_1 cms. and x_2 cms. with the radiator placed at the common centre in the equatorial plane, it can be shown readily that the ratio of the ionisation currents of excited and exciting radiations is given by

$$\frac{i_K}{i_\nu} = \frac{N_K}{n_\nu}\cdot\frac{\nu_K}{\nu_E}\cdot\frac{1-e^{-\tau_2(x_1-x_2)}}{1-e^{-\tau_1(x_1-x_2)}}\cdot e^{-\mu_1 x_1}\cdot e^{-\mu_2 x_2}, \qquad (7)$$

where $\mu_1 = \tau_1 + \sigma_1$, $\mu_2 = \tau_2 + \sigma_2$ are the absorption coefficients of the radiations ν_E and ν_K respectively in oxygen.

In the case of two plane electrodes parallel to the plane of the radiator, and distant from it x_1 and x_2 cms. respectively equation (7) becomes

$$\frac{i_K}{i_\nu} = \frac{N_K}{n_\nu}\cdot\frac{\nu_K}{\nu_E}\cdot\frac{1-e^{-\tau_2\psi(x_1-x_2)}}{1-e^{-\tau_1(x_1-x_2)}}\cdot e^{-\mu_1 x_1}\cdot e^{-\mu_2\psi x_2}, \qquad (8)$$

ψ is a factor which corrects for the increased length of oxygen traversed by the excited K rays corresponding to values of θ not zero.

The following assumptions were made in the derivation of equations (7) and (8). These assumptions are of particular interest in as much as they are

present in most measurements of X-ray intensities where ionisation methods are employed.

(*a*) The total radiation energy absorbed in producing ionised oxygen atoms (principally in the K shell) is employed ultimately in the production of " general " ionisation. This assumption implies that any radiations excited in the oxygen atoms is completely absorbed in the ionisation chamber. It will be shown later, however, that the efficiency of excitation of K series radiation becomes extremely small for atoms of low atomic number, and it seems probable that in the case of oxygen the radiation energy absorbed is completely converted into " photo-electron " energy.

(*b*) The energy required to produce a pair of ions is independent of the velocity of the ionising high speed photoelectron. On the basis of some early results of Lenard,* Holthusen† has proposed a correction to be used in the comparison of X-ray intensities when the beams are of different frequencies. According to Lenard's results the energy required to produce a pair of ions is a function of the velocity of the high speed ionising electron.

Considerable doubt has been thrown on the validity of this correction by some recent experiments of Kircher and Schmitz‡ and Kulenkampf,§ who found that the energy required to produce a pair of ions is independent of the frequency of the X-ray beam primarily responsible for the ionisation, at least in the wave length range 1·5 Å.U. to 0·5 Å.U. Further evidence for the probable constancy of the energy required to produce a pair of ions is discussed by Anslow.||

(*c*) Both the excited and exciting radiations produce high speed photo-electrons in the same length of oxygen, *i.e.*, $x_1 - x_2$ cms. This assumption is only justified when the ranges of the photoelectrons are small compared with the length of the chamber, or when the ranges associated with the excited and exciting radiations are nearly the same.

(*d*) The radiation scattered by oxygen makes no contribution to the ionisation currents. In the wave length range used the ionisation produced by Compton electrons can be neglected.

The sources of error (*c*) and (*d*) are discussed later in relation to the particular onisation chamber used.

* ' Quantitives über Kathodenstrahlen aller Geschwindigkeiten.'
† ' Fortschr. a. d. G. d. Röntgenstrahlen,' vol. 26, p. 215 (1919).
‡ ' Z. f. Physik,' vol. 36, p. 484 (1926).
§ ' Ann. d. Physik,' vol. 79, p. 97 (1926).
|| ' Phys. Rev.,' vol. 26, p. 484 (1925).

The values of the efficiency of K series emission are deduced finally from a combination of (6) and (7)

$$\left.\begin{aligned} f &= \frac{i_K}{i_\nu} \cdot \frac{\nu_E}{\nu_K} \cdot \frac{1}{A \cdot B}, \\ \text{where} \quad A &= 2 \Big/ \left[(1 - \cos \Omega) - \frac{\alpha_2}{\alpha_1} \log_e \frac{\alpha_1 + \alpha_2}{\alpha_1 \cos \Omega + \alpha_2} \right] \\ B &= 1 - e^{-\tau_1(x_1 - x_2)} / [1 - e^{-\tau_2(x_1 - x_2)}] \cdot e^{-\mu_1 x_1} \cdot e^{-\mu_2 x_2}. \end{aligned}\right\} \qquad (9)$$

§ 3. *Experimental Details.*

(*a*) *Source of Exciting X-Rays.*—As intense beams of X-rays of long wave length were required it was necessary to use an X-ray tube* provided with a thin aluminium window. This tube was of the hot cathode type connected continuously with a Crawford mercury jet pump (Western Electric type) and runs quite steadily when the current is not greater than 5 milli-amps.

A measure of the intensity of the X-ray beam issuing from the tube was obtained by placing a small ionisation chamber between the collimator slits of the spectrometer and measuring the ionisation produced with a Dolezalek electrometer. Since this ionisation chamber measures the total intensity of the beam it is important to keep the potential on the tube approximately constant, so maintaining constant the distribution of relative intensities among the various wave-lengths. To this end a voltmeter was placed across the primary of the transformer and the primary voltage controlled by means of variable resistances in series with the primary windings. The tungsten filament was heated from accumulators.

In order to obtain intense beams of exciting rays of wave lengths shorter than $Z_{n K_\alpha}$ it was found necessary to use the general radiation from a platinum target excited at appropriate voltages, and absorbed in layers of suitable salts. The process of filtering reduced the beam to at least 0·05 of the original intensity, and, judging from the curves obtained by plotting log I_0/I_d against thickness of aluminium, appears to yield beams of some degree of homogeneity. The absorbing materials used and the mean wave lengths of the filtered beams have been deduced from the aluminium absorption co-efficients obtained in the subsidiary experiment already referred to.

* The tube is described in detail elsewhere. Martin, ' Proc. Camb. Phil. Soc.' (*loc. cit.*).

Table II.

Filter.	Mean μ/ρ Al.	Mean λ.
$NH_4Br + 0{\cdot}125$ mm. Al	14·6	1·02 Å.U.
$Sr\,CO_3 + 0{\cdot}125$ mm. Al.	11·3	0·94 ,,
$Sr\,CO_3 + 0{\cdot}200$ mm. Al.	7·8	0·82 ,,
$Zr\,O_2 + 0{\cdot}400$ mm. Al	6·1	0·75 ,,
Mo $0{\cdot}133$ mm $+ 0{\cdot}800$ mm. Al	3·7	0·64 ,,

(*b*) *Apparatus.*—As the ionisation produced by the excited K radiations is only a small fraction of that produced by the exciting X radiation, a balance method was used in which it was possible to measure directly the ionisation produced by the excited rays while the total intensity of the exciting beam was deduced from the decrease in ionisation produced by absorbing the exciting beam in a thin aluminium foil of known thickness. The thickness of the aluminium was chosen so that the two ionisation currents measured were of the same order of magnitude. The method had the advantage that it enabled the Compton electrometer to be used continuously at the same sensitivity.

The apparatus used for measuring the absolute value of f for iron and nickel for homogeneous exciting radiations near the K limits is shown in fig. 1. It was used also for measuring the relative values of f for Fe Ni and Cu for the series of short wave length exciting radiations. The apparatus was mounted on an ordinary Bragg X-ray spectrometer in place of the usual ionisation chamber.

The exciting X-ray beam passes through the two chambers C_2 and C_1 by the windows W_2 and W_1 (holes 0·3 cm. in diameter) and falls on a radiator R which is supported on a small carbon disc. This disc can be rotated by means of an ebonite plug moving in a ground metal joint, and the radiator R turned out of the way of the beam of X-rays, which falls then on the carbon disc. The electrodes of this chamber (C_1) are hemispheres of radii 1·72 cms. and 5·05 cms. respectively. The radius of the inner hemisphere is greater than the maximum range in oxygen of the photoelectrons ejected from the radiator. This electrode was made of the thinnest paper available, gummed over a light framework of celluloid, and then painted with Indian ink to render it conducting. (The paper was so thin that it really formed a semi-transparent fine fibrous mesh.)

The outer hemisphere was painted very heavily with Indian ink in order to reduce as far as possible the emission of low speed electrons. The saturation ionisation currents were found to be independent of the sign of the voltage applied to the inner electrode within the limits of experimental error.

The balancing chamber C_2 is of the usual type with two plate electrodes.

A shield S can be moved by means of a milled nut N between the X-ray beam and the electrode connected with the electrometer. It was possible, by moving

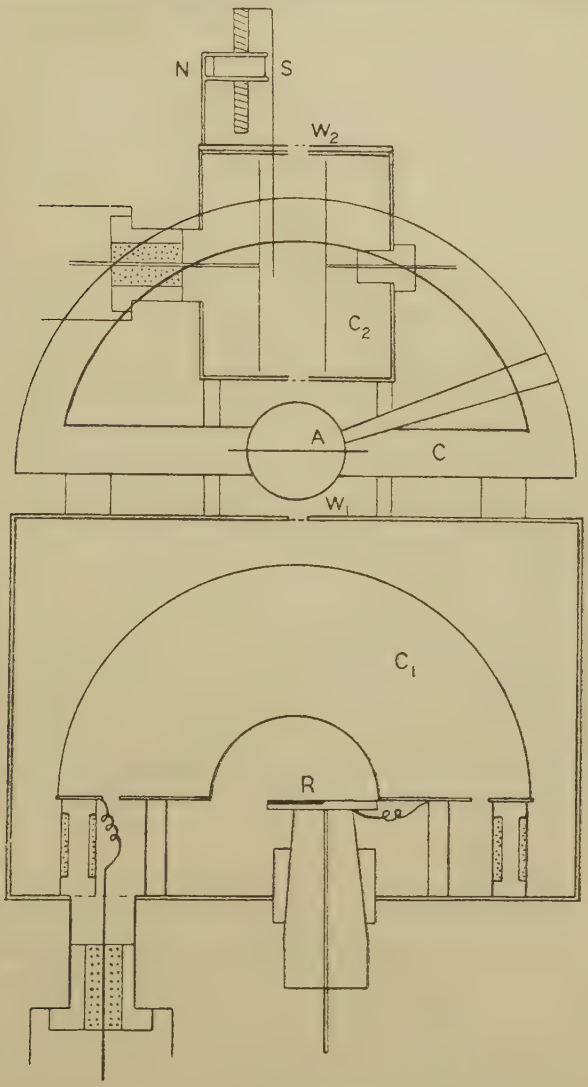


FIG. 1.

this shield to balance the ionisation current in C_1, and a balance once set up was found to hold through a run (usually 15 mins.).

The ionisation currents were measured by a Compton electrometer working at a maximum sensitivity of 8,000 mm. per volt. Small leaks were balanced out by applying a fraction of a volt to an artificial leak provided by a uranium oxide resistance.

For the softer exciting radiations, a difficulty arose in obtaining aluminium foils which were sufficiently uniform to justify a determination of their mass per unit area, and at the same time thin enough to produce a decrease in intensity in these radiations of only some 5 to 10 per cent. This difficulty was overcome by mounting an aluminium foil A (of thickness $D = 0{\cdot}00164$ cm.) on the moving arm of a divided circle C. If the aluminium foil is rotated through an angle θ, the decrease in intensity is equivalent to that caused by a foil of thickness

$$d = D\,[\sec\theta - 1].$$

(*c*) *Measurements.*—Determinations of the ratio of the intensity of the excited K radiations to the exciting radiations were carried out in the following manner. A balance between C_1 and C_2 was first obtained with an aluminium foil of thickness $d_{\text{cms.}}$ placed between the chambers and the radiator R turned out of the way of the exciting beam. The radiator is now turned into the path of the beam and the increased current "a" due to the excited K radiations is measured. The radiator is now turned out of the way, the aluminium foil removed and the current b measured. The ratio i_K/i_ν is then simply

$$\frac{i_K}{i_\nu} = \frac{k}{1-k} \cdot \frac{a}{b},$$

where k is the fractional decrease in the intensity of the exciting beam due to the interposition of $d_{\text{cms.}}$ of aluminium. In general, it is impossible to obtain a perfect balance and in the following table a and b are diminished by the balance current x. Table III shows typical sets of readings for two exciting beams $\lambda = 0{\cdot}637$ Å.U. $\lambda = 1{\cdot}537$ Å.U. respectively.

Table III.

Cu Radiator. Exciting radiation 0·637 Å.U. Al absorber $d = 0{\cdot}1356$ mm. $k = 0{\cdot}127$.

x Balance.	$a - x$ Excited K radiation.	$b - x$ $k/(1-k)$ exciting radiation.	i_K/i_ν $\frac{k}{1-k} \cdot \frac{a-x}{b-x}$.
7	42	40	0·146 . 1·050
9	42	42	1·000
8	42	38	1·105
9	48	38	1·262
11	44	42	1·048
8	50	44	1·137
8	46	46	1·000
10	46	44	1·036

Table III—(continued).

Fe Radiator. Exciting radiation Cu K_α 1·537 Å.U. Al absorber D = 0·00164 cm. $\theta = 50°$ $d = 0·000912$ cm. $k = 0·114$.

$\theta = 50°$ b	$\theta = 0°$ x	$\theta = 0°$ [Fe] a	i_K/i_ν $k\frac{a-x}{x-b}$.
0	36	59	0·114 . 0·638
−3	33	56	0·638
−4	34	59	0·658
0	37	59	0·595
1	37	59	0·612
1	35	55	0·588
−2	35	57	0·595

A slightly different procedure was carried out in the case of ZnK_α CuK_α and NiK_α, but this is sufficiently clear from the column headings. The deflections of the electrometer spot are always in the same direction and over the range used, some 100 mm., the electrometer sensitivity was found to be constant.

The values of i_K/i_ν have to be corrected first for the loss of intensity of the excited rays in passing through the inner electrode, secondly for the ionisation produced by electrons ejected from the outer electrode by the excited rays.

Measurements were made on the absorption by the celluloid frame of the K_α radiations of Fe, Ni and Cu, and the absorption deduced for the mean wave lengths of the K series radiations. The values for the fraction of radiation transmitted, after allowing for the variation of excited intensity with θ are

	Fe	Ni	Cu
Transmission Factor ..	0·63	0·70	0·73

A direct measurement in the case of the second correction showed that this amounted to 2 per cent. in the case of Fe K_α, 4 per cent. in the cases of Ni K_α and CuK_α.

The corrected values of i_K/i_ν and the deduced values of f are given in Table IV.

The following values were used in the evaluation of the factors A and B equation (9). Ω is defined by a copper ring which supports the inner electrode and has a value of 76° 10′. The values of α were taken from the measurements of the writer already referred to.

The values for the oxygen absorption coefficients have been calculated from the expression

$$\mu/\rho = 2·77\ \lambda^3 + 0·16.$$

Table IV.

Wave-Length Exciting rays.	—	i_K/i_ν.			Nos. of K quanta emitted per Ionised atom. f.		
		Fe.	Ni.	Cu.	Fe.	Ni.	Cu.
0·64 Å.U.	Rel. val.	1	1·184	1·386	1	1·17	1·31
0·75 ,, ...	,,	1	1·175	1·375	1	1·23	1·40
0·82 ,, ...	,,	1	1·160	1·393	1	1·25	1·46
0·94 ,, ...	,,	1	1·062	—	1	1·17	—
1·02 ,, ...	,,	1	1·069	—	1	1·16	—
1·433 ,, ...	Abs. val.	0·116	0·113	—	0·290	0·352	—
1·537 ,, ...	,,	0·108	—	—	0·293	—	—
1·655 ,, ...	,,	0·098	—	—	0·299	—	—
	Rel. val.	—	—	—	1	1·20	1·39

which represents a mean of the measurements of Hewlett,* Olsen Dersham and Storch,† and the general relation deduced by Richtmeyer and Warburton‡ for the atomic absorption coefficient. These experimenters have in every case made measurements for wave lengths which are considerably shorter than those which occur in the present experiment, but extrapolation according to a λ^3 law appears to be justified by the fact that a λ^3 law is found to hold in the case of a similar light element aluminium throughout the wave length range 0·6 Å.U. to 1·93 Å.U.

Absolute measurements were not possible with this type of chamber in the case of the shorter wave length exciting beams as the ranges of the photo-electrons produced by these beams are large compared with the length of the chamber. The use of the chamber for relative determinations of f is justified, however, as the constitution and intensity of these beams were kept constant.

In order to eliminate this source of error a new chamber, fig. 2, was constructed to replace C_1. In this chamber the plane electrodes are perpendicular to the direction of the X-ray beam, and one of them can be displaced through a known distance, 3·20 cms. If the ionisation currents are measured with the electrode first in one position, then in its displaced position, the difference after being corrected for the change in capacity, gives the ionisation currents produced by the absorption of the X-ray beam in a length of gas equal to the displacement. The chamber is lined with aluminium 0·6 mm. thick which in turn is covered with thick blotting paper. This eliminates any ionisation which would be pro-

* ' Phys. Rev.,' vol. 17, p. 284 (1921).
† ' Phys. Rev.,' vol. 21, p. 30 (1923).
‡ ' Phys. Rev.,' vol. 22, p. 329 (1923).

duced by the characteristic radiations of the brass chamber excited by scattered radiation from the exciting beam. This precaution is important as in the case

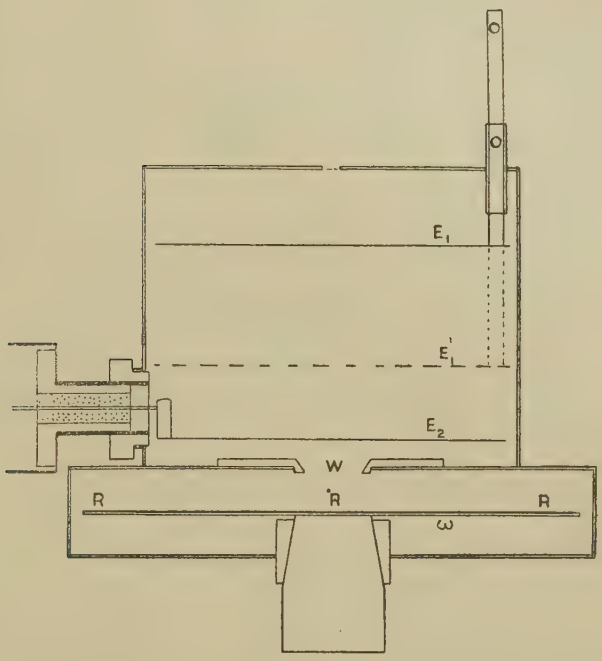


Fig. 2.

of radiation $\lambda = 0{\cdot}6$ Å.U. some 20 per cent. of the total radiation absorbed in oxygen is scattered.

The two electrodes E_1 and E_2 were made of very thin inked paper stretched tightly on an aluminium frame.

The exciting beam enters through a circular window 0·3 cm. diameter and passing through the centre of the window, W, falls on the radiators, R (Fe Ni Cu Zn Al), which are arranged round the circumference of a wheel *w*. The beam of excited radiation is limited by a thick aluminium circular window, W, which subtends an angle $2\ \Omega = 71° 8'$ at the surface of the radiators.

The experimental procedure is, in the main, similar to that of the previous experiments, but now the balancing chamber C_2 is replaced by a large uranium oxide leak. The capacity of this remains constant throughout the experiment; an ionisation current equal and opposite to that produced by the exciting beam in C_1 being obtained by exposing the requisite area of uranium oxide.

It was found necessary to use a very intense beam of exciting rays in order to obtain results of any degree of accuracy. This is due partly to the relatively

small value of Ω which can be used but the greatest source of error lies in the determination of the ionisation currents when E_1 is in its displaced position. The capacity is increased and the length of gas traversed only some 2 cms., with the result that the ionisation currents become quite small. An intense beam of radiation of high degree of homogeneity was obtained by filtering the radiation from a silver target (tube voltage 45 k.v.) through 0·05 mms. of palladium together with 0·6 mm. of aluminium. The mean wave-length was found to be 0·60 Å.U.

Typical results are shown in Table V. The first two columns give the ratio of excited to exciting intensities for each position of E_1; columns three and four the absolute values for the exciting beam for each position of E_1. These values are deduced as previously from the decrease in ionisation produced by interposing in the path of the beam an aluminium foil of known thickness. Column five gives the values for E_1 corrected for the capacity change.

Table V.—Excited K intensity.

Exciting intensity.		i_ν Exciting intensity.			i_K/i_ν.
E_1.	E_1'.	E_1.	E_1'.	E_1.	
0·0726	0·094	883	261	675	0·058
0·0763	0·120	852	250	650	0·050
0·0740	0·109	865	239	660	0·055
0·0776	0·125	890	200	680	0·057
0·0697	0·096	893	262	679	0·049

The final mean value of i_K/i_ν is 0·058. f is calculated from equations (6) and (8) which combined reduce to $i_K/i_\nu = f \times 0{\cdot}217$, giving a value for f equal to 0·268. Relative values of f were also measured for this wave-length.

——	Fe.	Ni.	Cu.	Zn.
i_K/i_ν	1	1·247	1·386	1·518
f	1	1·175	1·360	1·598

§ 4. *Results.*

The following table gives the collected results :—

Exciting Radiation. Wave-Length.	Relative values of f. Fe.	Ni.	Cu.	Zn.
0·64 Å.U. 1·433 Å.U.	1	$1·19_5$	$1·38_8$	—
0·603 Å	1	$1·17_5$	$1·36_0$	$1·59_8$
Mean values	1	1·19	1·38	1·60
	Absolute values of f.			
1·655 Å.U.	0·299	—	—	—
1·537 ,, ...	0·293	—	—	—
1·433 ,, ...	0·290	0·352	—	—
0·603 ,, ...	0·268	—	—	—
Weighted mean value	0·285	0·352	—	—
Deduced absolute values	0·29	0·35	0·40	0·46

The values for the efficiency ϕ of K emission in K ionised atoms are deduced from the values of f by the relation

$$\phi = f \cdot \frac{J_K}{J_K - 1}$$

where J_K is the K jump. It was found, however, in the case of Fe Ni and Cu that the exponent of λ in the law of variation of μ/ρ with λ had different values on each side of the K limit,* *i.e.*,

$$\lambda > \lambda_K \quad \mu/\rho = \text{const. } \lambda^{2·98}$$
$$\lambda < \lambda_K \quad \mu/\rho = \text{const. } \lambda^{2·80}.$$

It has been found, however, that for shorter wave-lengths than those considered here ($\lambda < 0·6$ Å.U.) a similar relation holds on each side of the K discontinuity.† As a result the K jump for one of these elements increases gradually from its value at the K limit, and approaches a limiting value in the neighbourhood of $\lambda = 0·6$ Å.U. The following are the limiting values found for J_K :—

K jump.

	Allen.	Martin.
Fe	9·5 → 10	9·2 → 10
Ni	8·8 → 9·8	8·5 → 9·6
Cu	9·1 → 9·8	8·6 → 10
Zn	8·8 → 9·5	—

* Allen, ' Phys. Rev.' (*loc. cit.*) ; Martin, ' Proc. Camb. Phil. Soc.' (*loc. cit.*).

† Allen, ' Phys. Rev.,' vol. 26, p. 266 (1926) ; Richtmeyer and Warburton, ' Phys. Rev.,' vol. 22, p. 539 (1923).

Efficiency of K emission for K ionised atoms.

——	Fe.	Ni.	Cu.	Zn.
K Limit	$0{\cdot}32_4$	$0{\cdot}39_5$	$0{\cdot}44_9$	$0{\cdot}51_9$
0·6 Å.U.	$0{\cdot}32_2$	$0{\cdot}38_9$	$0{\cdot}45_1$	$0{\cdot}51_4$

It was pointed out earlier that it was possible to deduce the values of ϕ for bromine selenium and iodine from some early experiments of Barkla and Beatty on the ionisation produced in these gases by a series of homogeneous X-radiations. These measurements are tabulated in Table VI.

Table VI.

Exciting Radiations.		Ionisation relative to air.		
K series.	K_α.	SeH_2.	C_2H_5Br.	CH_3I.
Fe	1·932 Å.U.	x	y	—
Ni	1·655 ,,	x	y	162
Cu	1·537 ,,	x	y	152
As	1·170 ,,	x	y	158
Br	1·041 ,,	x	y	—
Sr	0·871 ,,	$4{\cdot}1\ x$	$3{\cdot}4\ y$	—
Mo	0·708 ,,	$6{\cdot}3\ x$	$4{\cdot}4\ y$	188
Ag	0·562 ,,	$7{\cdot}7\ x$	$5{\cdot}4\ y$	198
Sn	0·487 ,,	$8{\cdot}3\ x$	$5{\cdot}9\ y$	205
I	0·437 ,,	—	—	211
Ba	0·388 ,,	—	—	251 [β_{12}= 365]

If there are n quanta of radiation of frequency ν passing into the ionisation chamber the number absorbed in a layer δx of gas distant x from the entrance is

$$n\, e^{-\mu x}\, (\mu_K + \mu_L)\, \delta x$$

where $\mu = \mu_K + \mu_L$ is the total absorption coefficient. The ionisation produced by photoelectrons ejected from the L level is

$$\text{const.}\ n\, e^{-\mu x} \,.\, \mu_L \,.\, \delta x \,.\, h\,(\nu - \nu_L),$$

while that produced by the K photoelectrons is

$$\text{const.}\ n\, e^{-\mu x} \,.\, \mu_K \,.\, \delta x\, [\phi h\,(\nu - \nu_K) + (1 - \phi)\, h\nu]$$

where ϕ is the efficiency of K emission for K ionised atoms. It has been assumed that with the ionisation chambers used the ionisation produced by the absorption of excited K rays in the gas is negligible.

The total ionisation current produced in a chamber of length d is

$$\text{const. } n h \left[\nu\mu - \phi\, \nu_K\, \mu_K\right] \int_0^d e^{-\mu x}\, \delta x.$$

On the long wave length side of a K discontinuity, the ratio of the ionisations x, y is

$$\frac{c_1}{c_2} \cdot \frac{1 - e^{-{}_1\mu d}}{1 - e^{-{}_2\mu d}}$$

while on the short wave-length side the ratio is

$$\frac{c_1}{c_2} \cdot \frac{\mu_2}{\mu_1} \cdot \frac{1 - e^{-\mu_1 d}}{1 - e^{-\mu_2 d}} \cdot \frac{\mu_1\nu - \phi\nu_K\mu_K}{\mu_2\nu}$$

or

$$\frac{c_1}{c_2} \cdot \frac{1 - e^{-\mu_1 d}}{1 - e^{-\mu_2 d}} \left[1 - \phi \frac{\nu_K}{\nu} \cdot \frac{J_K - 1}{J_K}\right]$$

where J_K is the K absorption jump. At the K limit $\nu_K = \nu$, $\mu_1 = J_K \cdot {}_1\mu$, ${}_2\mu = \mu_2$ and the ratio of the ionisations increases by an amount R given by

$$R = J_K \left[1 - \phi \frac{J_K - 1}{J_K}\right]. \qquad (10)$$

The values assigned to R are deduced by extrapolating the values tabulated in Table VI to the respective K limits (fig. 3). The following values were found for R :—

	Se	Br	I
R	3·1	3·1	1·7

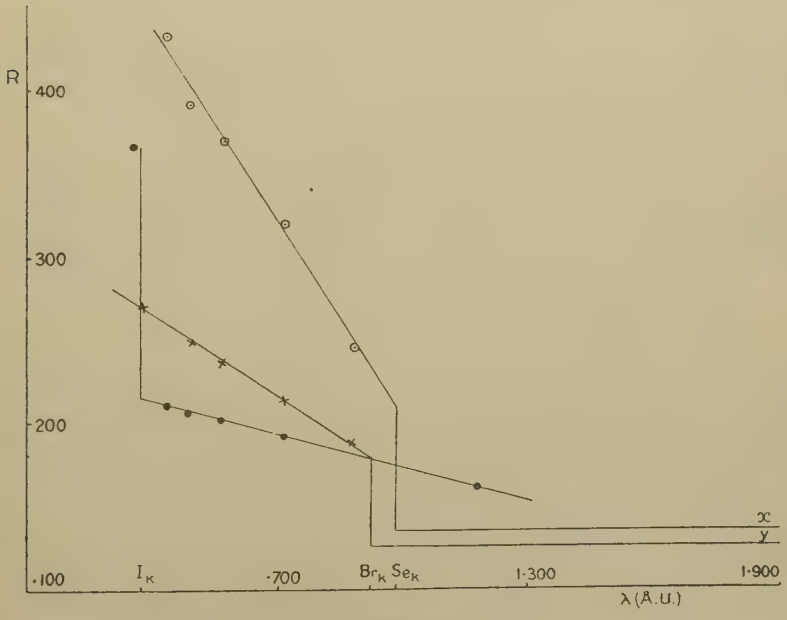


FIG. 3.

Actual determinations of the K jump have only been made in the case of $C_2 H_5 Br$. This measurement was made by Barkla and Sadler using radiations which were not strictly homogeneous, and it was thought better to use values deduced from determinations made under better conditions for neighbouring elements.

	Se	Br	I
J_K	7·7	7·7	6·8

Substituting these values of R and J_K in equation (10) the following values are found for ϕ :—

	Se	Br	I
ϕ	0·68	0·68	0·88

It is interesting to notice that, if the value of J_K found by Barkla for Br is used, *i.e.*, $J_K = 5·3$, $\phi = 0·51$. This value of ϕ was deduced by Bothe (*loc. cit.*) from the results of Barkla and Thomas.*

If $\phi = 0$, *i.e.*, if no characteristic radiation is emitted from an ionised atom the curve obtained by plotting R against the wave-length will be horizontal. This is the case for Br and Se on the long wave length side of their K limits (fig. 3) indicating that for these atoms all the L series radiation is internally converted. This is not the case, however, for iodine. Unfortunately, as it is impossible to evaluate c_1/c_2, ϕ_L cannot be deduced from this curve. Auger has found in the case of the neighbouring element Xenon $\phi_L = 0·25$.

§ 5. *Discussion of Results.*

(*a*) *Variation of ϕ_K with atomic number.*—The different values of f_K assigned by different observers are collected in Table VII, and values of ϕ are shown plotted against atomic number in fig. 4.

Table VII.—Efficiency of K Emission.

Atom. Nos.	Kossel.	Auger.	Bothe.	Balderston.	Martin.
18	—	0·07	—	—	—
24	0·23	—	—	—	—
26	0·32	—	0·25	0·33	0·29
27	0·39	—	↑	—	—
28	—	—	↓	0·39	0·35
29	0·42	—	0·25	0·44	0·40
30	0·51	—	—	0·50	0·46
34	—	—	—	—	0·59
35	—	—	0·50	—	0·59
36	—	0·51	—	—	—
42	—	—	—	0·83	—
47	—	—	0·93	0·75	—
53	—	—	—	—	0·75
54	—	0·70	—	—	—

* Barkla, ' Bakerian Lecture ' (*loc. cit.*).

The values given by Kossel were deduced from the application of Holthusen's* correction to the results of Barkla and Sadler (Table I). It has already been

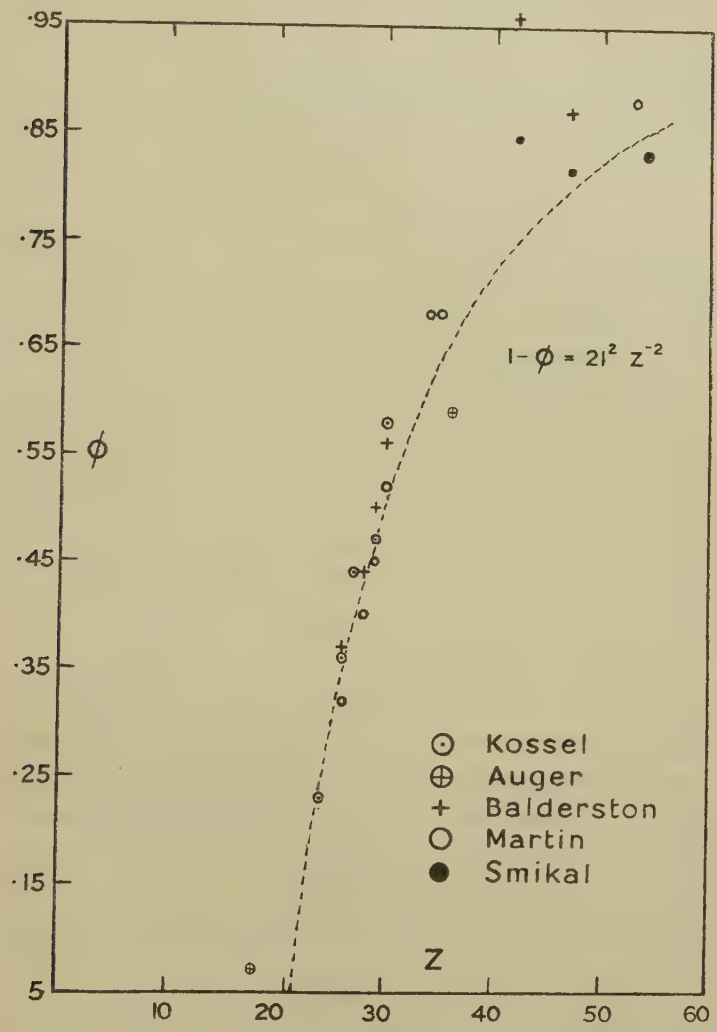


FIG. 4.

mentioned that there is considerable doubt as to the validity of this correction. The values obtained here are in good agreement with those obtained by Barkla and Sadler for exciting wave lengths near the K limits, but are consistently lower than the values given by Balderston. The error in the present experiments is in the neighbourhood of 5 per cent., and it is probable that the sum of the

* Holthusen (*loc. cit.*).

mean errors of the two experiments is at least as great as the observed differences in ϕ some 10 per cent.

The values of f deduced for Se Br and I fit in well with the determinations of Balderston for Mo and Ag, and the value assigned to Ag by Bothe from some results by Barkla and Dallas on the photoelectron emission from this element. These values are far removed however from the curve of the empirical relation suggested by Smikal,* *i.e.*,

$$1 - f = 16 \cdot 2/(Z - 1), \tag{11}$$

which represents satisfactorily the results of Auger and the collective results for the Fe group.

This empirical relation obviously fails to represent values of ϕ for elements of atomic number lower than 17, but a closer investigation of the variation of ϕ with atomic number would lack point until more data are available for elements of higher atomic number. In particular it will be important to determine whether the rapid decrease observed by Balderston in ϕ for Ag is real and is maintained for heavier elements, or whether ϕ gradually approaches the limiting value 1. It is interesting to notice, however, that the general variation of ϕ with Z is well represented by the following alternative equation—

$$1 - \phi = 21^2 \, Z^{-2}. \tag{12}$$

This is shown in fig. 4 as a dotted line.

This relation shows a striking similarity to that representing the normal absorption of K series radiation. To a sufficient approximation the fraction F of the energy of a beam of X-rays of 1 cm.^{2} section which is absorbed by the L M, etc., groups of electrons of any atom is given by the well-known relation

$$F = \text{const. } Z^4 \lambda^3.$$

Since the wave length λ of the K series radiation is approximately proportional to Z^{-2},

$$F = \text{const. } Z^{-2}. \tag{13}$$

This striking similarity between the laws which represent the variation with atomic number of normal absorption and "internal" absorption cannot be stressed however, as rather large differences exist in the values of ϕ for elements of high atomic number. The single determination of ϕ made by Auger for an element of very low atomic number (*i.e.*, Z = 18) indicates that the variation of $1 - \phi$ with Z is no longer given by a relation similar to (12), but one in which

* 'Ann. d. Physik,' vol. 81, p. 391 (1926). A critical discussion is given here of the available data referring to "radiationless" phenomena in radioactive and X-ray regions.

the exponent of Z is positive. This difficulty may not necessarily destroy the significance of the similarity pointed out above as there is some evidence* that the normal absorption increases at a much slower rate than λ^3 in the case of elements of very low atomic number. From the data at present available the general similarity existing between the laws of " normal " and " internal " X-ray absorption is favourable to the general picture of these radiationless reorganisations which is suggested by the " internal " absorption process.

(*b*) *Variation of ϕ with Wave-length of Exciting Radiation.*—Although the value of ϕ_{Fe} found for $\lambda = 0{\cdot}603$ Å.U. is some 7 per cent. smaller than the mean of the values for $\lambda = 1{\cdot}433$ Å.U., $1{\cdot}537$ Å.U. and $1{\cdot}655$ Å.U. respectively, the experimental error in the case of the first is of this order, and it must be concluded that within these limits ϕ is independent of the wave length of the exciting beam. A similar result was found by Balderston in the wave-length range $0{\cdot}306$ Å.U. $\rightarrow 0{\cdot}887$ Å.U., and it has now been shown in the case of Fe, at least, that the efficiency of K series emission is independent of the wave length of the ionising radiation from the K limit to $0{\cdot}306$ Å.U. This leads to the natural conclusion that the " internal conversion process " is a mechanism to be associated with the ionised atom itself, and is quite independent of the radiation field responsible for the ionised state of the atom.

It appears that this independence of frequency applies over a much wider range of frequencies than has been investigated in the X-ray region. Some important evidence on this point has appeared recently in the experiments of Ellis and Wooster† on the relative intensities of the natural β-ray lines in the spectra of radium B and radium C. Certain pairs of lines are produced by the internal conversion of the same γ-ray in the L and K limits, and it is found that the ratio of the intensities of these lines is $0{\cdot}22$ over the range of γ-rays investigated ($5{\cdot}07 \times 10^{-10}$ cms.$\rightarrow 0{\cdot}866 \times 10^{-10}$ cms.). It is most significant that the value of this ratio is that which would be obtained for lead on the assumption that the ordinary laws of absorption for X-rays can be extrapolated to this region.

(*c*) *Energy Balance of K series Transformations.*—Although the agreement shown between the different values of ϕ for elements of high atomic number leaves much to be desired, a closer agreement could hardly be hoped for as some of the experiments from which values of ϕ have been deduced were not devised to that particular end. It will be seen from fig. 4, however, that the results do form a consistent and coherent whole in spite of large individual variations.

* Holweck, ' C. R.,' vol. 171, p. 849 (1920); ' C. R.,' vol. 172, p. 439 (1921).

† ' Roy. Soc. Proc.,' A, vol. 114, p. 276 (1927).

In some experiments the energy of the K characteristic radiation excited is measured, in others the energy of the photoelectrons ejected by the " internal absorption " of the K radiation, and as these complementary phenomena lead to values of ϕ, which agree among themselves within the limits of experimental error, it may be assumed with some confidence that the energy exchanges and transformations associated with the K ionisation of an atom and its subsequent reorganisation are now accounted for quantitatively.

Summary.

The K series characteristic radiations have been excited in a series of radiators Fe, Ni, Cu and Zn, by different exciting radiations (0·6 Å.U. → K limit) and the efficiency of K emission determined by an ionisation method.

The low values found for this efficiency together with the results of other related experiments find an explanation in the hypothesis that in some cases owing to conditions as yet undetermined the excited K radiation does not escape the atom, but is " internally absorbed " in outer electron shells giving rise to high speed electrons.

It is found that the probability of K series emission as opposed to its internal absorption is independent of the frequency of the exciting radiation within the limits of experimental error.

The efficiency of K series emission is a function of the atomic number of the excited atom, and is such that it seems probable that similar laws, relating the probability of absorption with atomic number and wave-length, hold in the cases of "normal" and " internal " absorption.

In conclusion, the author wishes to express his thanks to Sir Ernest Rutherford for his constant interest and advice throughout this work, and acknowledge also his indebtedness to Dr. Chadwick and Dr. C. D. Ellis for helpful criticism and discussion.

Currents Carried by Point-Discharges beneath Thunderclouds and Showers.

By T. W. Wormell, B.A., St. John's College, Observer in Meteorological Physics at the Solar Physics Observatory, Cambridge.

(Communicated by C. T. R. Wilson, F.R.S.—Received May 6, 1927.)

[Plate 7.]

1.—*Introduction.*

The present paper describes experimental methods employed for investigating the sign and magnitude of the vertical current due to the point-discharge of electricity from an elevated metal point, during the intense electric fields associated with thunderstorms and showers. The results of preliminary observations extending over a period of about eight months are tabulated. It is found that there is a considerable net transfer of positive electricity upwards from the earth to the atmosphere by the discharge currents. The vertical discharge current is thus in the opposite direction to the normal fine-weather ionization current and to the convection current carried by various forms of precipitation. A comparison of the magnitude of the observed quantities with that of the fine-weather current indicates the important part played by the point-discharge currents in the total exchange of electricity between the earth and the atmosphere. Finally, the distribution of electric field below a cumulo-nimbus cloud deduced from the observed variations in magnitude and direction of the current, as the cloud passes overhead, strongly suggests that the great majority of shower-clouds and thunder-clouds observed were bipolar clouds of positive polarity, *i.e.*, the upper charge was positive.

2.—*The Discharging System.*

The discharging apparatus is carried at the top of a vertical wooden pole, whose base is attached by a hinge to a stout wooden stake, fixed vertically in the ground. The pole can thus be lowered to the ground for inspection of the insulation. When the pole is in its vertical position, the height of the discharging point above the surface of the ground is 8·3 metres. The lead carrying the current to the measuring apparatus consists of about 30 metres of lead-covered rubber insulated cable, of external diameter 5·6 mm. The

cable travels down the pole and thence, underground, to the hut containing the observing apparatus. The distance from the base of the pole to the hut is about 14 metres. Any disturbance of the field at the top of the pole, due to the presence of the hut is negligible.

Fig. 1.

The discharging point is represented in fig. 1. It consists of a 4 cm. length of fine copper wire, A, of diameter 0·25 mm., attached to the top of a vertical stout iron wire B, of length 30 cm. The lower end of B is clamped by a small screw, C, to the closed top of a brass tube D. The dimensions of D are :—length 4 cm., diameter 1·75 cm. The tube D with its attached wires, slides on to a solid cylindrical piece of brass, E, which is cemented to the top of a quartz tube H, of length 8 cm. The wire from the upper end of the cable leading to the measuring apparatus passes through the quartz tube H, and is clamped by a screw to E. The lower end of the quartz tube is fixed in contact with the end of the rubber insulation of the cable, inside the upper end of the brass tube K by sealing wax. This joint is strengthened by a wrapping of insulating tape (not shown in the figure). The base of the exposed portion of the quartz tube H and the upper portion of the brass tube K, are liberally coated with vaseline to prevent insects from crawling upwards on to the quartz insulation. The brass "umbrella," D, protects the quartz tube from rain and makes quantitative observations possible during heavy rainfall. When each end of the cable has been protected by a quartz tube, as illustrated for the upper end in fig. 1, the insulation of the whole system is quite satisfactory; it is regularly tested as a matter of routine and is usually found to be unaffected by rainfall.

If q be the charge carried by the discharging system and the lead to the hut, the value of $-1/q\ dq/dt$, *i.e.*, that fraction of the charge lost per second, which is due simply to leakage across the insulation, is normally of the order of 1/700. This corresponds to a leakage resistance of about 10^{11} ohms. On one or two isolated occasions a much greater leak was detected, due to the presence of spider-threads across the quartz at the top of the pole. The current due to the point discharge from the apparatus described, has been investigated by two entirely independent methods.

3.—*Methods of Observation.*

(*a*) *Integration of the current by a voltameter.*—The first method employed consists in integrating the current during a period of intense electric field, by means of a special form of gas microvoltameter, which is connected between the end of the cable from the discharging point, and earth. The method of the gas voltameter has the advantage that, when the volumes of gas liberated from both electrodes are separately measured, the quantity of electricity which has passed in each direction may be deduced. Thus, if 1 unit of electricity liberates c c.c. of oxygen and $2c$ c.c. of hydrogen, then, if v_1 c.c. of gas are liberated from electrode A and v_2 c.c. from electrode B during a definite period, and if a quantity q_1 of electricity enters the voltameter at A and leaves at B, and a quantity q_2 enters at B and leaves at A during that period, we have :—

$$cq_1 + 2cq_2 = v_1,$$
$$2cq_1 + cq_2 = v_2.$$

From these equations we have at once

$$q_1 - q_2 = (v_2 - v_1)/c,$$
$$q_1 + q_2 = (v_1 + v_2)/3c.$$

$q_1 - q_2$ is the net quantity which has passed in one direction ; and $q_1 + q_2$ the total quantity which has passed in both directions.

The actual design of the voltameter is shown in figs. 2 and 3. Fig. 2 is a

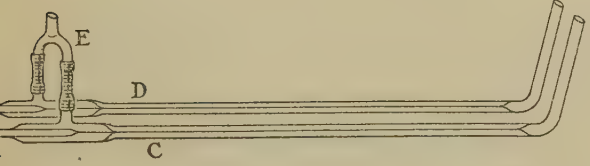


Fig. 2.

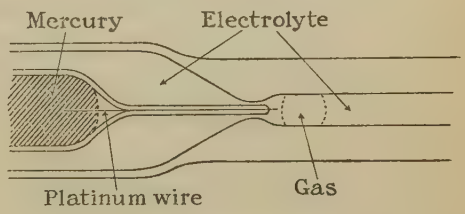


Fig. 3.

sketch of the complete apparatus. In fig. 3 a portion of one electrode (near C in fig. 2) is drawn on a larger scale. The two electrodes A and B are sealed into wider glass tubes, and project into the ends of two horizontal capillary tubes C and D. The circuit is completed through the tube E, the shaded portions being rubber pressure tubing. A slight constriction is formed at the near end of each capillary tube, the electrode projecting in each case just beyond the narrowest part of the constriction. When electrolysis is proceeding, the main bubble of gas collects with its near end at the point where the diameter

of the capillary tube becomes constant. The electrodes consist of platinum wire of 0·025 mm. diameter passing down fine thin walled capillary tubes, through the ends of which they are sealed. The wire projects into the electrolyte for a length of about 0·1 mm. The other end of the platinum wire is immersed in mercury, whence connections may be made with an external circuit. The electrolyte is dilute sulphuric acid of strength 1 in 10. The internal diameter of the capillary tubes C and D is about 0·8 mm. and the passage of 10^{-2} coulomb gives a bubble of hydrogen of length about 2·5 mm. The total length of the voltameter is 30 cm. and the height to the top of E, 10 cm. When the voltameter is in regular use, it is convenient to minimise evaporation of the liquid by connecting the tops of the three open tubes together by rubber tubing and a suitable T-piece.

The design adopted for this instrument was arrived at in attempting to avoid, as far as possible, various sources of error which might become important. In the first place, in dealing with small volumes of gas, of the order of a cubic millimetre, the possibility of solution of the gas in the electrolyte must be reduced to a minimum. The gas is collected in the horizontal capillary tubes at a distance of about 1 mm. from the electrode, and thus the volume of liquid to be saturated with gas is small. Again, during a thunderstorm, the sign of the electric field and hence the direction of the current may be frequently reversed by lightning flashes. Under such circumstances, bubbles of mixed hydrogen and oxygen will be collected. It was found that if such a bubble remained in contact with the platinum electrode, recombination of the gases was so rapid that the motion of the boundaries of the bubble, as it decreased in size, was visible with a low power microscope. The introduction of a constriction at the end of the tube, as already described, causes the bubble to move a short distance away from the electrode. No appreciable recombination has been detected in a mixed bubble when not in contact with platinum. There is, however, a slow decrease in the size of a bubble, whether it be pure hydrogen or oxygen or a mixture of the two. This effect is presumably due to slow solution of the gas in the electrolyte. For bubbles of a few mm. in length, the rate of decrease in length is rather less than 0·2 mm. per day. Another source of error, if frequent current reversals occurred, might reside in the capacity of the electrodes. The area of platinum exposed to the electrolyte has, however, been made so small that this effect is quite negligible in the form of apparatus described, the capacity of the electrodes being of the order of 2×10^{-3} microfarad. No current will pass through the instrument unless the difference of potential across the electrodes

exceeds about 2 volts. For the present purpose this is an advantage, since any accidental electromotive force in the circuit will not produce a spurious current provided the value of the electromotive force remains below 2 volts. The resistance of the voltameter is of the order of 10,000 ohms. Thus, for a current as large as a milliampere, the potential difference between its terminals would only be of the order of 12 volts. The effect of this voltage on the magnitude of the discharge-current is entirely negligible, since the electric field would be of the order of 10,000 volts per metre or more.

A voltameter of the form described is a convenient apparatus for measuring quantities of electricity from about 10^{-2} coulomb upwards, with an accuracy of about 10^{-3} coulomb. Within this accuracy, the volume of gas liberated does not seem to depend on the previous history of the electrodes nor on the magnitude of the current, for currents from 0·1 microampere up to 200 microamperes.

(*b*) *Observations by a spark gap and capillary electrometer.*—In the second method, observations have been made on the actual values attained by the current in the discharger system during a period of intense electric field. The main difficulty in such observations lies in the enormous range over which the magnitude of the current may vary. As a storm approaches, the magnitude of this current may increase in a few seconds from something undetectable by the methods to be described (a value less than, say, 10^{-10} ampere) to 10^{-5} ampere or more, a range of values of 1 to 10^5. The method of observation adopted has the advantage that measurements are possible over a very wide range of values of the current, and that no damage will be done to the apparatus by currents of unexpected magnitudes. The method readily permits, moreover, of obtaining a permanent photographic record of the variation of the current with time.

The arrangement of the apparatus is represented diagrammatically in fig. 4. A spark gap S is inserted between the end of the lead from the discharger, and earth. When a point-discharge current is flowing, the large capacity C_1 of the lead is charged up until its potential reaches the value V at which a spark passes across the gap, when the potential falls very rapidly almost to zero. The passage of a spark is considered for the present purpose, to reduce the potential instantaneously to zero. In order to measure the potential difference at which a spark passes across the gap, and also to obtain when desired a record of the variation of the current through the discharger with time, a small air condenser C_2 and a capillary electrometer E are connected in series, across the spark gap. The capillary electrometer is of the form developed

by Prof. C. T. R. Wilson* and used by him in his researches on the electric fields of thunderstorms. It measures by the displacement of the meniscus the quantity of electricity which has passed through it, and maintains one plate of the condenser C_2 at zero potential. By a suitable recording device a trace of the variation of the position of the meniscus with time is obtained on a photographic plate. The recording arrangement is described in the reference; the source of illumination is now a small electric lamp, supplied with current from storage cells. The condenser C_2 consists of two parallel circular brass plates carried on quartz rods clamped in metal supports on an earthed base, which prevents any leakage current through the condenser and electrometer. The plates of this condenser are 5 cm. in diameter, the distance apart being usually about 3 mm. The capacity is thus of the order of 5 cm. The capacity C_1 of the lead is 5200 cm. A spark gap of about 0·3 mm. is usually employed. In order to provide a rapid means of determining the direction of the current when visual observations are being made, a neon lamp of the ordinary commercial " Osglim " type is inserted at N. The appearance of the flash in the lamp as each spark occurs affords a convenient and quite definite indication of the direction of the current. The lamp appears to affect the regularity of the sparking to a certain extent, and, consequently, when accurate results are desired, it is short circuited by the key K. The direction of the current is then determined by watching the motion of the meniscus of the electrometer or by recording that motion photographically. The motion of the meniscus consists of a slow creep as the capacity C_1 charges up and a sudden jump back to its zero position when a spark passes. The displacement of the meniscus at any instant, from its zero, is a measure of the charge at that instant on one plate of the condenser C_2. The magnitude of the sudden jump when a spark passes measures the charge q on one plate of

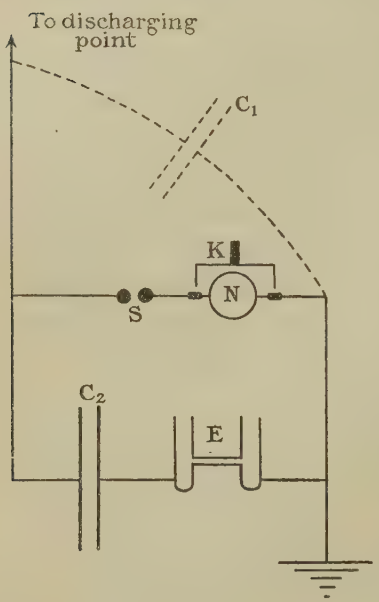


Fig. 4.

* ' Roy. Soc. Proc.,' A, vol. 92, p. 555 (1916); and ' Phil. Trans.,' A, vol. 221, p. 74 (1920).

this condenser, when the potential difference, V, between the plates is equal to the sparking potential at the gap. We have—

$$q = C_2V.$$

The quantity of electricity passing to or from earth at each spark is C_1V and if N is the frequency of the sparks, the current i is given by the equation—

$$i = NC_1V = N(C_1/C_2)\, q.$$

The integrated current is given at once by the total number of sparks. By visual observations, values of N ranging from one in several minutes up to a frequency giving a note of definite pitch could presumably be measured. When a photographic record is being obtained, however, the present apparatus will not record higher rates of sparking than about 3 per second.

The sparking method imposes a rather severe strain on the insulation of the system. As has been previously stated, the fraction of the charge lost per second due to insulation losses is usually of the order of 1/700. Thus the error due to this cause is about 1 per cent., when sparks are occurring at the rate of 1 per 14 secs. The rates of sparking usually observed are of the order of 1 per second while the centre of a shower is overhead. Slower rates necessarily occur during the beginning and end of the shower. The error in the deduced value of the current, due to insulation leakage, is quite negligible except when the sparks are occurring less frequently than one per 10 seconds. The error due to this cause, in estimating the integrated current during a typical shower is probably less than 1 per cent. The accuracy is further limited by the precision with which q, the quantity discharged by each spark may be determined, and is estimated to be rather better than 5 per cent.

If the plates of the condenser C_2 be short circuited, the discharging apparatus is connected through the electrometer, to earth, and the whole current passes through the electrometer. With this arrangement, visual observations can conveniently be made on the sign and magnitude of currents ranging from about 10^{-10} to 10^{-8} ampere. Considerable caution is necessary in using this method, as the electrometer may readily be electrolysed and put out of action by too large a current.

4.—*The quantities of electricity discharged from the apparatus.*

The values obtained for the quantities of electricity transferred between the earth and the atmosphere by the artificial discharging point already described, during various showers and thunderstorms, are tabulated below. The method of observation employed in each case is indicated in the second

column. "a" indicates that the voltameter was used, "b" that observations were made with the sparking apparatus. In the next three columns are tabulated:—q_1, the quantity of electricity passing from the earth to the atmosphere; q_2, the quantity passing from the atmosphere to the earth, and $q_1 - q_2$, the net loss of positive electricity by the earth. All these quantities are in millicoulombs. The observations with the voltameter are accurate to about 1 millicoulomb, those with the spark gap to about 0·2 millicoulomb. In the last column, notes on the accompanying meteorological conditions are given in some cases. The showers referred to in this column are, in general, phenomena of fairly brief duration, lasting frequently about 30 minutes.

Date.	Method.	q_1.	q_2.	$q_1 - q_2$.	——
1925. September and October	a	—	—	40	Rough observations on the effect of several heavy showers accompanied in two cases by thunder.
December 15	a	11	2	9	Fall of snow of about 25 mm. depth.
1926. February 16	a	14	1	13	A short thunderstorm. Two lightning flashes were within 1 km. of the apparatus.
April 17	a	13	0	13	A short thunder-shower. Three flashes were observed at about 6 km. distance.
,, 18	a	5	7	−2	Prolonged steady rain.
,, 21	a	17	2	15	A thunderstorm and several heavy showers during the night.
,, 22	a	6	0	6	Moderate rainfall with occasional heavy showers.
,, 24	a	0	0	0	Persistent gentle rainfall.
,, 27–28	a	2	2	0	Moderate rainfall with heavier showers.
,, 30	a	0	0	0	Fairly heavy rain.
May 7	a	—	trace	−trace	Gentle rainfall.
,, 12	a	11	0	11	Several showers.
,, 13	a	30	0	30	Heavy shower with thunder.
,, 20	a	9	9	0	Several hours of steady rainfall.
,, 21	a	3	4	−1	Distant thunderstorm followed by steady rain for hours.
	b	4	1	3	Dark cumulus cloud forming on a sultry afternoon. No appreciable rainfall.
,, 26	a	2	0	2	—
,, 30	a	4	0	4	Two heavy showers.
,, 31	a	0·5	0·5	0	—
June 1	b	1·2	0·0	1·2	Moderate shower.
,, 2	a	3	0	3	Steady rain for 24 hours with occasional heavy showers.
,, 9	a	1	1	0	Gentle rain and a heavy shower.
,, 10	b	5·0	1·2	3·8	A series of thunder showers.
,, 11	a	2	2	0	Showers with thunder.
,, 12–16	a	0·5	0	0·5	Overcast sky, much gentle rain and some heavy showers.
,, 17	a	1	5	−4	—
,, 19	a	1	0	1	Heavy cumulus clouds and only a trace of rain.

Date.	Method.	q_1.	q_2.	q_1-q_2.	——
June 24	a	1	5	−4	The extreme edge, only, of a large shower cloud passed overhead.
,, 27	a	1	1	0	Heavy shower.
July 5	a	2	2	0	Several hours of steady rain.
,, 18–19	a	11	11	0	Severe thunderstorms during the night. Flashes were occurring at the rate of about 15 per minute at 1 a.m. Nearest distance of approach was about 4 or 5 km.
,, 21	a	3	1	2	Slight shower in early afternoon.
	a	3	0	3	Very heavy shower with thunder. One flash was about 0·5 km. distant.
,, 25	a	0·5	0	0·5	Moderate showers.
,, 27	a	2	3	−1	Slight rainfall in the evening. Distant thunderstorms had been observed during the day.
August 7–30	a	(14)	(0)	(14)	This was approximately the total quantity indicated by the voltameter after a three weeks' absence. It includes several thunder showers.
September 1	a	3	5	−2	Very heavy shower. Several flashes were observed at a distance of 5 km.
	a	3	2	1	Heavy shower.
,, 2–7	a	0	0	0	Much rain and several heavy showers.
,, 23	a	4	0	4	Thunder shower. The nearest flash was about 2 km. distant.
,, 24	a	trace	—	trace	Very heavy shower.
,, 25	a	1	4	−3	Shower. Several flashes but not within 10 km.
	a	4	1	3	Shower.
,, 27	a	1	1	0	Moderate rainfall.
	a	2	0	2	Further rain after a fair interval.
October 8	a	trace	—	trace	Heavy rain and gale.
,, 10	a	1	0	1	Slight showers.
	b	4·0	0·2	3·8	Shower.
,, 12 and 15	a	0	0	0	Prolonged heavy rain.
,, 21	a	3	3	0	—
,, 24	a	5	0	5	—
,, 31	a	0·5	—	0·5	—
November 5	a	trace	—	trace	Heavy shower.
,, 8	a	10	2	8	Exceptionally heavy shower.
	a	8	0	8	Moderate rainfall.
,, 9	b	10	3	7	Heavy shower.
,, 13–14....	a	10	0	10	Heavy showers and persistent rain.
,, 20	a	2	0	2	Persistent rain and drizzle.
,, 26	a	trace	trace	—	Showers.
December 23	a	trace	trace	—	Drizzle.
,, 25	a	trace	—	trace	Slight rainfall.
Total		255	82	173	

We see that, during the showers and thunderstorms indicated in this table, a quantity $0\cdot17_3$ coulombs of positive electricity was transferred from the earth to the atmosphere, due simply to the current from a single discharging point. While the observations are not entirely complete, they include the majority of occasions during the months April to December, 1926, on which the electric

field was sufficiently intense to cause an appreciable point discharge current, *i.e.*, of the order of 1 microampere, for a period of 10 minutes or more. It would seem that the figure $0{\cdot}17_3$ coulombs would indicate the order of magnitude of the total effect due to the point discharge current from the apparatus, in stormy weather during these months. It is of interest to compare the observed quantity with the transfer of electricity during the same period due to the normal fine weather ionization current. If we take a value 2×10^{-12} ampere per sq. metre as an average value of this fine weather current, we find that the observed quantity of electricity discharged from the point would just neutralize the fine weather current into an area of ground of about 3700 sq. metres during the same period. The results can be expressed in a slightly different way. It will be seen from the table that it was frequently found that, during a heavy shower, lasting perhaps, 30 minutes, a net quantity of positive electricity of the order of 10 millicoulombs was discharged from the earth by the single point under observation. Such a quantity will neutralize the fine weather current into an area of about 160 sq. metres during a whole year. These figures will, perhaps, suffice to indicate the importance of the effect investigated. It is clear that in any discussion of the total transfer of electricity between the earth and the atmosphere, the effect of the point-discharge currents from prominent natural objects, such as trees, during periods of intense electric field, cannot be neglected. One would expect the results obtained with the apparatus described to be of the same order of magnitude as the quantity of electricity discharged, under the same conditions, from a small tree.

During the observations described, no really severe thunderstorm occurred within about 4 km. of the observing apparatus.

5. *Observations on the actual values of the current and their bearing on the electrical structure of cumulo-nimbus clouds.*

Some observations by means of the sparking apparatus will now be considered in more detail. The particular cases selected for consideration have been chosen because of the completeness of the information given by a single record.

Fig. 5, Plate 7, is a record of the vertical current during a brief shower on June 1, 1926. The values of the vertical potential gradient are also recorded just before, and just after, the shower. The record of this potential gradient was obtained by measuring the charge on an exposed earth-connected metal sphere

by means of a capillary electrometer.* The phenomena recorded were produced by an isolated small shower cloud of cumulo-nimbus type, which appeared to pass right overhead. No other shower clouds were in the immediate vicinity at the time. At 12 h. 30 m. the approach of the cloud was noticed. At 12 h. 37 m. there was a small positive potential gradient of + 25 volts per metre. This gradually increased to a value of about + 50 volts per metre, then decreased to zero and soon reached large negative values. The potential gradient was zero at about 12 h. 39 m. At 12 h. 40 m. it had reached the value — 470 volts per metre and at about 12 h. 41 m. — 890 volts per metre. At about this time fairly heavy rain began to fall and the metal sphere was lowered to the ground and shielded. The discharge current became sufficiently large to cause sparks to pass across the gap at about the same time. From this point onwards the photograph records the number of sparks passing across the gap and the values of the vertical current at each moment may be deduced. The quantity of electricity passing across the gap at each spark was 4·5 microcoulombs. The current was upwards throughout the shower, corresponding to a negative potential gradient. The current attained a maximum value of 3·5 microamperes at 12 h. 44 m. The total quantity of electricity discharged from the discharging point during the shower was $1{\cdot}2_3 \times 10^{-3}$ coulombs. The current became inappreciable at about 12 h. 53 m. Just before this, the potential gradient must have been negative and exceeded, say, — 500 volts per metre. At 12 h. 57 m. 10 s. the sphere was again raised and the potential gradient was found to be positive with a value of 350 volts per metre. The value of the potential gradient was decreasing and at 12 h. 59 m. it had reached a value 200 volts per metre. The values of the potential gradient and vertical current deduced from this plate have been plotted in the figure. No thunder was heard during this shower and there is no indication of sudden changes in the current such as would be produced by lightning discharges.

Fig. 6, Plate 7, is an enlargement of part of a record of the current obtained on June 10, 1926. It illustrates the detailed information which can be obtained by this method. On this occasion a series of clouds of cumulo-nimbus type was passing overhead, and in several of the clouds lightning discharges were occurring. The variations in the value of the current are thus due to the superposition of two causes, the motion of the charged clouds relative to the apparatus, and variations in the magnitudes of the charges due to lightning flashes and other causes. Sudden increases of upward current occur at

* C. T. R. Wilson, *loc. cit.*

12 h. 13 m. 33 s. and 12 h. 16 m. 13 s., indicating sudden increases of negative potential-gradient due to lightning discharges. The distances of these two discharges, deduced from the interval between the discharge and the moment when the thunder was first heard, are 3·1 km. and 2·1 km. respectively. The time of occurrence of the thunder is recorded by momentarily interrupting the illumination, and hence producing dark lines on the record, TT. At 12 h. 19 m. 31 s. and 12 h. 20 m. 55 s. sudden positive changes of potential gradient clearly occurred, due to discharges at distances of 4·9 km. and 4·7 km. respectively. After each of these four sudden changes, the variation of the current suggests that the electric-field showed a typical "recovery curve"* after a lightning discharge.

The type of distribution of electric field below a shower-cloud indicated in fig. 5 is similar to that suggested by visual observations of the discharge current during a considerable number of showers. The observations are supplemented in some cases by measurements of the electric field before and after the fall of rain. The phenomena frequently noticed as a shower-cloud approaches and passes overhead are that the observed current is first downwards, then in the middle of the shower, upwards, and finally, towards the end of the shower, downwards again. There is, in such cases, a large negative potential gradient beneath the central portion of the cloud, and surrounding this area, smaller positive potential gradients which may, however, be considerably greater than the normal potential gradient of fine weather. In some cases, it is only in the intense negative field beneath the centre of the cloud that a measurable discharge current occurs. Such was the case in the shower studied in fig. 5, but the measurements of the field in this case indicate that the distribution of electric field was of the type under consideration.

The tabulated observations of the quantities of electricity discharged during showers and thunderstorms also support this suggestion as to the typical distribution of electric field. In the majority of cases there is a net transference of electricity upwards from the earth to the atmosphere during the storm or shower. Some of the occasions on which the net transfer of electricity was zero or, in the opposite direction, can clearly be interpreted by assuming that the centre of the storm did not pass overhead. In some cases, when a downward transfer of electricity was observed, it could be seen that only the fringe of a shower-cloud passed overhead. It has not, however, been shown that all cases of a net downward transfer of electricity can be explained in this way.

A distribution of electric field of the type indicated would be produced*

* C. T. R. Wilson, *loc. cit.*

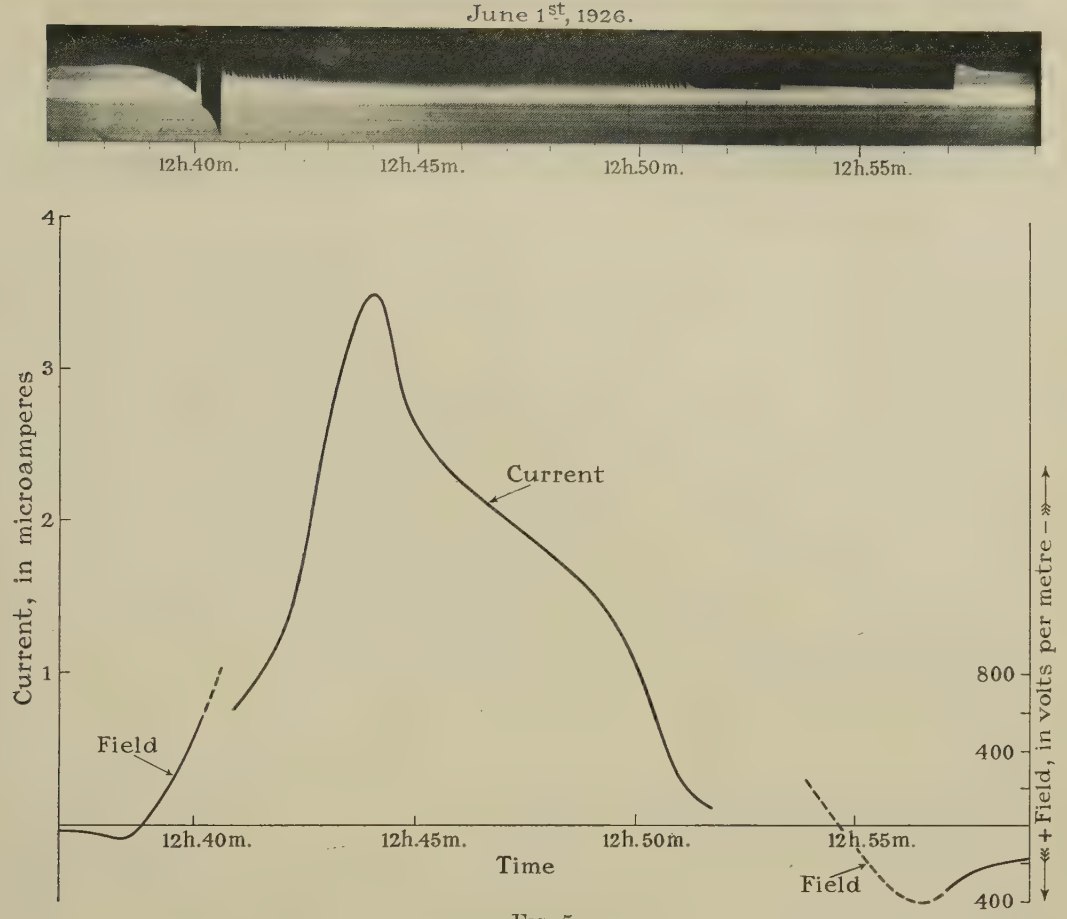


FIG. 5.

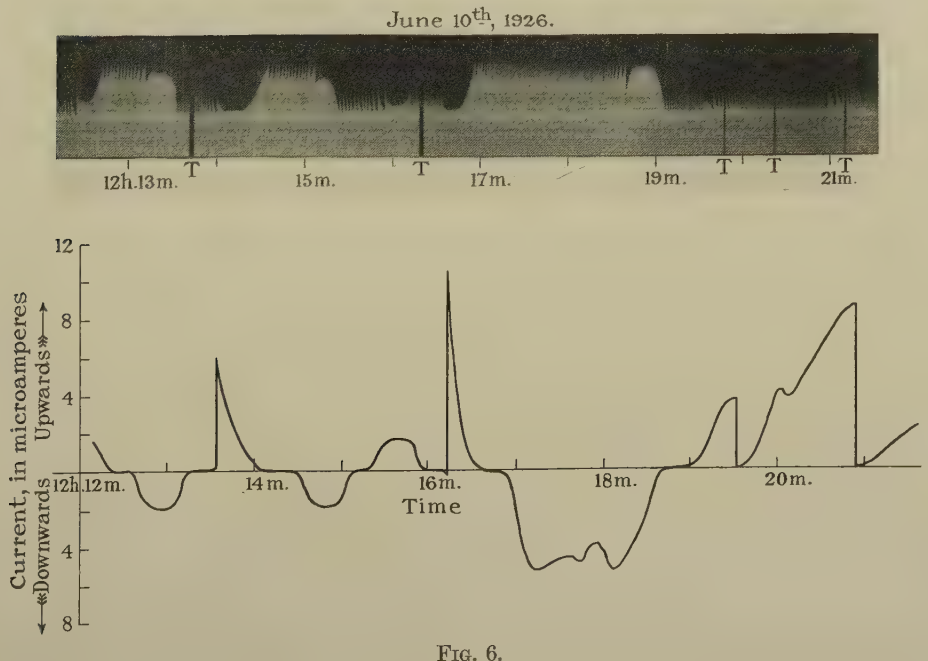


FIG. 6.

(*Facing p.* 454)

under certain conditions by a positive bipolar cloud, *i.e.*, a cloud of which the upper charge was positive, the lower charge negative. In a recent paper, Schonland and Craib* describe observations on the electric fields of thunderstorms in South Africa and interpret their results as indicating that, in practically all the cases observed, the clouds were of positive polarity. Appleton, Watt and Herd† combining their own observations with previous ones by Prof. Wilson on the effects of lightning flashes on the electric-field in this country, also conclude that thunderclouds must frequently be of positive polarity. In the present series of observations on the discharge-current, there appears to be no great difference in the orders of magnitude of the effects due to small thunderstorms and those due to some showers in which no lightning discharges occur. The observations strongly suggest that, in the majority of cases, the cumulo-nimbus cloud is of positive polarity. It is, perhaps, unnecessary to mention the importance of this conclusion in theories of the method of maintenance of the earth's negative charge.

In conclusion, it gives me great pleasure to acknowledge my indebtedness to Prof. C. T. R. Wilson, who suggested the observations described and the experimental methods employed. I have had the benefit of his continual interest and advice. My thanks are also due to the Department of Scientific and Industrial Research for a grant.

The observations described in this paper were carried out at the Solar Physics Observatory, Cambridge. I am very grateful to the Director for the facilities accorded me before I became a member of the staff. Some of the apparatus was constructed and tested at the Cavendish Laboratory.

Observations by the methods described are being continued.

* 'Roy. Soc. Proc.,' A, vol. 114, p. 229 (1927).
† 'Roy. Soc. Proc.,' A, vol. 111, p. 654 (1926).

The Crystal Structure of α-Manganese.

By A. J. Bradley, M.Sc., Ph.D., and J. Thewlis, M.Sc., the Physical Laboratories, the University of Manchester.

(Communicated by W. L. Bragg, F.R.S.—Received May 6, 1927.)

I. *Introductory.*

It has been shown independently by Westgren and Phragmén,* and by one of the present authors,† that manganese is allotropic. Both investigations pointed to the existence of three crystalline modifications. The names given to the different forms of the metal were not the same in the two papers, and in order to avoid ambiguity we have adopted Westgren and Phragmén's nomenclature. This latter is preferable because the so-called electrolytic manganese, which we called α-manganese (their γ-manganese) is probably a hydride of the metal. It therefore seems better to give the names α-manganese and β-manganese to the two forms which are known to be metallic, as W. and P. have done. The present paper describes the structure of α-manganese, which is stable at room temperature and higher temperatures up to about 700° C.

In the two investigations referred to above, the powder method of X-ray analysis alone was employed. In order that the results of any method of analysis may be trustworthy, the method must be capable of discriminating between all solutions which are theoretically possible, and it is not sufficient merely to indicate one solution which explains the observations. If the experimental data are not extensive, and the accuracy with which they can be tested is low, it is easy to get a spurious agreement between observation and calculation, which lies well within the limits of experimental error. This is especially the case with a structure such as that of α-manganese where the unit cell contains a large number of atoms, and the atomic positions depend upon a large number of parameters.

It is possible to avoid ambiguity either by increasing the amount of experimental data or by increasing the accuracy of the measurements. In the present case we were limited to results obtained by means of the powder method of analysis, because it has not yet been possible to get single crystals of α-manganese which could be used for rotation photographs or for an investigation with the spectrometer. It was therefore necessary to work with powder-photographs

* 'Z. f. Physik,' vol. 33, p. 777 (1925).

† A. J. Bradley, 'Phil. Mag.,' vol. 50, p. 1018 (1925).

which were so accurate that the classification of the lines in the photograph could be carried out without the possibility of error. This necessitated a much higher precision of measurement than has usually been obtained in the employment of the powder method.

The precision camera designed by Phragmén and used by him and Westgren in their investigations on manganese* seems admirably suited for this purpose, being a great improvement on any previous design. The principle involved in its construction is essentially that of the Bohlin camera. An accuracy and resolving power are attained which establish beyond doubt the dimensions of the crystal lattice and limit the space group to very few possibilities. We have therefore utilised the experimental data published by Westgren and Phragmén in the course of the present investigations. After a long process of elimination we have arrived at a solution for the complete atomic arrangement which we think is the only one capable of explaining the data.

2. *Size of Unit Cell and Number of Atoms.*

Westgren and Phragmén's measurements lead without ambiguity to a cubic cell whose side is 8·894 Å. In their original paper they assigned 56 atoms to each unit cell. In the development of our analysis we were led to question the accuracy of this number (n). It appeared that the relative intensities of the different lines could be explained much more satisfactorily by supposing there were 58 atoms in the unit cell, instead of 56. This correction, first suggested solely by the analysis, has proved to be justified by a redetermination of n. It is by no means so easy a point to decide as would appear at first sight because pure α-manganese is so porous that density determinations are not reliable. This porosity is partly but not completely avoided by using impure specimens.

Westgren and Phragmén determined the value of n from a specimen containing 0·4 per cent Al, 1·3 per cent. Fe and 1·5 per cent. Si. They have kindly informed us of an arithmetical error in their calculation of the mean atomic weight of the specimen which led them to get too low a value for n. In their paper they give 56·33. A recalculation of their data shows that they should have got 57, but this figure is not reliable because the specimen contained both α- and β-manganese. We have recently redetermined the value of n from a specimen of the composition given above but alloyed with 0·35 per cent. C, which turns the whole material into α-manganese. The density was found to be 7·26. The mean atomic weight is 54·17. The lattice dimensions of the specimen are $a = 8{\cdot}903$ Å. The number of atoms per unit cell is there-

* *Loc. cit.*

fore 57·32. The next paragraph shows that the structure is body centred and therefore has an even number of atoms per unit cell. This supports our conclusion that there are 58 atoms per unit cell, and that the somewhat lower figure obtained experimentally must be ascribed to the influence of porosity on the density determination.

3. *Possible Space Groups.*

The data given in Tables I and II show that planes with $(h + k + l)$ odd have their spacings halved, all other spacings being normal. There are therefore only a limited number of space groups which are possible. An inspection of Astbury and Yardley's* tables shows which space groups fulfil these conditions. They all correspond to body centred cubic lattices, and the possible space groups are T^3, T_h^5, T_d^3, O^5, O_h^9.

Summing up the results so far obtained, we may conclude that α-manganese is based on a body centred cubic lattice the space group being T^3, T_h^5, T_d^3, O^5 or O_h^9, the side of the cube being 8·894 Å, and each cell containing 58 atoms.

4. *Experimental Intensity Data.*

The following tables give the available information relating to the relative intensities of the lines. Table I was obtained from a powder photogram of the usual type. The intensity data of Tables II, III, IV are identical with those given by Westgren and Phragmén in their paper on manganese, but are slightly rearranged to distinguish the results obtained with the different cameras which they employed in their investigation.

The distinction between results obtained with different cameras is essential since the intensities of lines from different films are not directly comparable. A line described as "13" in their paper is omitted as it has not been observed in later photographs and is therefore spurious. The three cameras from which the data of Tables II, III and IV were obtained are precision cameras of the Bohlin type designed by Phragmén, each camera covering a limited range of reflexion angles. The figures in the last column of each table are calculated for the structure which we propose for α-manganese, and should be correlated with the intensity of the observed lines. These figures are calculated from the empirical formula $NS^2/(h^2 + k^2 + l^2)$, N being the frequency factor, giving the number of planes corresponding to $(h\,k\,l)$ and S the structure amplitudes. The latter was calculated from the factors given below. We have used this simplified intensity formula, omitting some of the terms

* 'Phil. Trans.,' vol. 224, p. 221 (1924).

Table I.—Powder Photogram of α-Manganese from Westgren and Phragmén's Observations.

$h^2 + k^2 + l^2$.	$h\,k\,l$.	$NS^2/(h^2 + k^2 + l^2)$.	Observed Intensity.
2	110	0·02	absent.
4	200	0·06	absent.
6	211	0·32	absent.
8	220	0·70	absent.
10	310	0·10	absent.
12	222	0·58	absent.
14	321	0·53	absent.
16	400	3·78	weak.
18	411 300	89·4	very strong.
20	420	0·40	absent.
22	322	27·6	strong.
24	422	9·4	medium.
26	510 431	13·9	medium.
30	521	1·10	weak.
32	440	0·25	absent.
34	433 530	1·14	very weak.
36	600 442	1·38	weak.
38	532 611	2·32	weak.
40	620	0·37	absent.
42	541	0·38	absent.
44	622	3·4	weak.
46	631	0·73	absent.
48	444	8·6	strong.
50	550 710 543	11·6	strong.
52	640	0·01	absent.
54	721 633 552	47·9	very strong.
56	642	7·5	medium.
58	730	5·4	medium.

Table II.—Precision Photogram of α-Manganese from Westgren and Phragmén's Observations—*First Camera.*

$h^2 + k^2 + l^2$.	$h\,k\,l$.	$NS^2/(h^2 + k^2 + l^2)$.	Observed Intensity.
6	211	0·32	weak.
8	220	0·70	weak.
10	310	0·10	very weak.
12	222	0·58	weak.
14	321	0·53	weak.

Table III.—Precision Photogram of α-Manganese from Westgren and Phragmén's Observations—*Second Camera.*

$h^2 + k^2 + l^2$.	$h\,k\,l$.	$NS^2/(h^2 + k^2 + l^2)$.	Observed Intensity.
16	400	3·78	weak.
18	411 330	89·4	strong.
20	420	0·40	absent.
22	322	27·6	strong.
24	422	9·4	medium.
26	510 431	13·9	strong.
30	521	1·10	weak.
32	440	0·25	absent.
34	433 530	1·14	very weak.
36	600 442	1·38	weak.
38	532 611	2·32	weak.

Table IV.—Precision Photogram of α-Manganese from Westgren and Phragmén's Observations—*Third Camera.*

$h^2 + k^2 + l^2$.	$h\,k\,l$.	$NS^2/(h^2 + k^2 + l^2)$.	Observed Intensity.	
			α_1.	α_2.
34	433 530	1·14	weak	—
36	600 442	1·38	weak	weak.
38	532 611	2·32	weak	very weak.
40	620	0·37	absent	absent.
42	541	0·38	absent	absent.
44	622	3·4	medium	weak.
46	631	0·73	weak	very weak.
48	444	8·6	strong	medium.
50	550 710 543	11·6	strong	medium.
52	640	0·01	absent	absent.
54	721 633 552	47·9	very strong	strong.
56	642	7·5	strong	medium.
58	730	5·4	medium	weak.

usually inserted, because it seems scarcely possible in this case to get a formula which is really sound theoretically. Such being the case it seems best to use the simplest possible empirical formula. The general agreement

between calculation and observation constitutes the evidence for the correctness of the analysis which will be described in the following paragraphs.

5. *Possible Atomic Groupings.*

In whatever manner the atoms are arranged they must certainly satisfy all the requirements of T^3 since each of the possible space groups includes among its conditions all the conditions imposed by T^3. We shall show in section 6 that there is only one arrangement of 58 atoms which satisfied the demands of T^3 and accounts for the observed intensities of reflexion. This is therefore the only structure which is in agreement with the experimental data.

To satisfy the requirements of T^3, atoms may be distributed in any of the following ways :—

Positions with No Degrees of Freedom :—

X $(0\,0\,0)$ $(\frac{1}{2}\,\frac{1}{2}\,\frac{1}{2})$ 2 atoms.

Y $(\frac{1}{2}\,\frac{1}{2}\,0)$ $(\frac{1}{2}\,0\,\frac{1}{2})$ $(0\,\frac{1}{2}\,\frac{1}{2})$ $(\frac{1}{2}\,0\,0)$ $(0\,\frac{1}{2}\,0)$ $(0\,0\,\frac{1}{2})$ 6 atoms.

Positions with One Degree of Freedom :—

A $(a\,a\,a)$ $(a\,\bar{a}\,\bar{a})$ $(\bar{a}\,a\,\bar{a})$ $(\bar{a}\,\bar{a}\,a)$ $(\frac{1}{2}+a, \frac{1}{2}+a, \frac{1}{2}+a)$ $(\frac{1}{2}+a, \frac{1}{2}-a, \frac{1}{2}-a)$
$(\frac{1}{2}-a, \frac{1}{2}+a, \frac{1}{2}-a)$ $(\frac{1}{2}-a, \frac{1}{2}-a, \frac{1}{2}+a)$ 8 atoms.

B $(b, \frac{1}{2}, 0)$ $(\bar{b}, \frac{1}{2}, 0)$ $(0, b, \frac{1}{2})$ $(0, \bar{b}, \frac{1}{2})$ $(\frac{1}{2}, 0, b)$ $(\frac{1}{2}, 0, \bar{b})$ $(\frac{1}{2}+b, 0, \frac{1}{2})$ $(\frac{1}{2}-b, 0, \frac{1}{2})$
$(\frac{1}{2}, \frac{1}{2}+b, 0)$ $(\frac{1}{2}, \frac{1}{2}-b, 0)$ $(0, \frac{1}{2}, \frac{1}{2}+b)$ $(0, \frac{1}{2}, \frac{1}{2}-b)$ 12 atoms.

C $(c, 0, 0)$ $(\bar{c}, 0, 0)$ $(0, c, 0)$ $(0, \bar{c}, 0)$ $(0, 0, c)$ $(0, 0, \bar{c})$
$(\frac{1}{2}+c, \frac{1}{2}, \frac{1}{2})$ $(\frac{1}{2}-c, \frac{1}{2}, \frac{1}{2})$ $(\frac{1}{2}, \frac{1}{2}+c, \frac{1}{2})$ $(\frac{1}{2}, \frac{1}{2}-c, \frac{1}{2})$ $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}+c)$ $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}-c)$
12 atoms.

Positions with Three Degrees of Freedom :—

D (x, y, z) $(x, \overline{y}, \overline{z})$ $(\overline{x}, y, \overline{z})$ $(\overline{x}, \overline{y}, z)$
$(\frac{1}{2}+x, \frac{1}{2}+y, \frac{1}{2}+z)$ $(\frac{1}{2}+x, \frac{1}{2}-y, \frac{1}{2}-z)$ $(\frac{1}{2}-x, \frac{1}{2}+y, \frac{1}{2}-z)$ $(\frac{1}{2}-x, \frac{1}{2}-y, \frac{1}{2}+z)$
(y, z, x) $(\overline{y}, \overline{z}, x)$ $(y, \overline{z}, \overline{x})$ $(\overline{y}, z, \overline{x})$
$(\frac{1}{2}+y, \frac{1}{2}+z, \frac{1}{2}+x)$ $(\frac{1}{2}-y, \frac{1}{2}-z, \frac{1}{2}+x)$ $(\frac{1}{2}+y, \frac{1}{2}-z, \frac{1}{2}-x)$ $(\frac{1}{2}-y, \frac{1}{2}+z, \frac{1}{2}-x)$
(z, x, y) $(\overline{z}, x, \overline{y})$ $(\overline{z}, \overline{x}, y)$ $(z, \overline{x}, \overline{y})$
$(\frac{1}{2}+z, \frac{1}{2}+x, \frac{1}{2}+y)$ $(\frac{1}{2}-z, \frac{1}{2}+x, \frac{1}{2}-y)$ $(\frac{1}{2}-z, \frac{1}{2}-x, \frac{1}{2}+y)$ $(\frac{1}{2}+z, \frac{1}{2}-x, \frac{1}{2}-y)$
24 atoms.

These groups of atoms may be combined in various ways, subject to the condition that the total number in the unit cell is 58. Trial shows that the atoms may theoretically be arranged in any one of the following 21 ways. Which

of these corresponds with reality is decided later from a consideration of the relative intensities of reflexion.

(1) $8A_1$, $8A_2$, $8A_3$, $8A_4$, $8A_5$, $8A_6$, $8A_7$, 2X.
(2) 12C. } $8A_1$, $8A_2$, $8A_3$, $8A_4$, $8A_5$, 6Y.
(3) 12B. }
(4) $12B_1$, $12B_2$ }
(5) 12B, 12C } $8A_1$, $8A_2$, $8A_3$, $8A_4$, 2X.
(6) $12C_1$, $12C_2$ }
(7) $12B_1$, $12B_2$, $12B_3$ }
(8) $12B_1$, $12B_2$, 12C } $8A_1$, $8A_2$, 6Y.
(9) 12B, $12C_1$, $12C_2$ }
(10) $12C_1$, $12C_2$, $12C_3$ }
(11) $12B_1$, $12B_2$, $12B_3$, $12B_4$ }
(12) $12B_1$, $12B_2$, $12B_3$, 12C }
(13) $12B_1$, $12B_2$, $12C_1$, $12C_2$ } 8A, 2X.
(14) $12B_1$, $12C_1$, $12C_2$, $12C_3$ }
(15) $12C_1$, $12C_2$, $12C_3$, $12C_4$ }
(16) 24D, $12B_1$, $12B_2$, 8A, 2X.
(17) 24D, 12B, 12C, 8A, 2X.
(18) 24D, $12C_1$, $12C_2$, 8A, 2X.
(19) 24D, 12B, $8A_1$, $8A_2$, 6Y.
(20) 24D, 12C, $8A_1$, $8A_2$, 6Y.
(21) $24D_1$, $24D_2$, 8A, 2X.

The structure amplitudes of the various groupings are given by the moduli of the complex quantities ($i = \sqrt{-1}$), as follows :—

X. $S = +2$.

Y. $S = +6$ for h, k, l all even.
$S = -2$ for planes with mixed indices.

A. $S = 8 \cos ha \cos ka \cos la - 8i \sin ha \sin ka \sin la$.

B. $S = 4 \cos hb \cos \frac{k}{2} + 4 \cos kb \cos \frac{l}{2} + 4 \cos lb \cos \frac{h}{2}$.

C. $S = 4 \cos hc + 4 \cos kc + 4 \cos lc$.

D. $S = 8 \cos hx \cos ky \cos lz + 8 \cos hy \cos kz \cos lx$
$+ 8 \cos hz \cos kx \cos ly - 8i \sin hx \sin ky \sin lz$
$- 8i \sin hy \sin kz \sin lx - 8i \sin hz \sin kx \sin ly$.

6. *The Selection of the Correct Structure.*

In this paragraph we describe the method of elimination by which we were led to select one of the above alternative structures as being the only one capable of explaining the observed results. The process is lengthy, since so many possibilities are involved. It is described in some detail to show that the solution is unique. Each case is considered in turn with the following results.

Cases 1–15.—Any atomic arrangement *not involving D atoms* must give large values either to all reflexions from (h 0 0) planes or to all from (h h 0) planes. An examination of the structure factors for the groupings X, Y, A, B, C will show clearly that this must be the case. As a matter of fact (110), (200), (220) and (440) are absent from the powder photogram and are either absent or weak on the precision photograms. These arrangements are therefore impossible.

Cases 16–20.—The principle underlying the analysis in these cases is as follows. The nature of the structure amplitude formulæ for the X, Y, A, B or C atoms is such that unless the parameters have very particular values, quite appreciable reflexions will be observed from all planes with indices (h 0 0) or (h h 0). Most of these reflexions are observed to be very weak, and we have tried to find parameter values which would explain this while at the same time accounting for the very large reflexion from (330) and (411). This is, however, quite impossible for cases 16–20, as will be seen from some of the figures given.

The calculations are made more lengthy by the presence in each case of a set of D atoms. These are capable of three independent parameter variations. The effect of changing one parameter depends so largely on the values of the other two parameters, that it is necessary to consider every possible value of one parameter simultaneously with every possible value of the other two parameters. In order to treat the problem in a practical way, we have confined our attention to a finite number of parameter values. The co-ordinates of a typical atom were arbitrarily chosen so as to be of the form $l/16, m/16, n/16$, where l, m, n are integers. There are 4096 such positions in the unit cell, but symmetry operations make them very largely identical. The number of parameter values investigated in this way was large enough to ensure that for planes for which Σh^2 was small, no great change in structure amplitude occurred when passing from one of the above positions to a neighbouring position. We were therefore sure that the approximate positions of the atoms would be indicated by a rough agreement between the observed and calculated intensities for several slightly different parameter values. If no such agreement could be found it could be taken for granted that the structure being tested was incorrect.

The first test applied to the above parameter values was the following. Which of them, if any, could account for the very strong reflexions from (330) and (411) ? Calculation showed that no possible contribution from the A, B, C, X or Y atoms alone was at all sufficient to account for such a strong line. It was apparent that the D atoms were supplying quite a large proportion of the reflected radiation. We therefore calculated the contributions given by the D atoms to these planes for each of the selected parameter values. It was found that only 15 independent parameter values gave any appreciable contribution to either (330) or (411). These 15 sets of values are collected in Column I of Table V and are numbered I–XV. Any of the arrangements 17–20 would require the D atoms to have one of these fifteen sets of parameter values. No values of the D parameters can satisfy arrangement 16.

The second test was then applied. Which of these fifteen sets of parameter values, if any, could explain the non-occurrence of the reflexions from (110), (200), (220), (310), (420) and (440) ? Each of the five arrangements 16–20 was tested separately.

Table V.—Value of S for D Atoms with Various Parameter Values.

Typical D atom.	411.	330.	110.	200.	220.	420.
i. $\frac{7}{16}$ $\frac{7}{16}$ 0	$+6\cdot8+i\,(0)$*	$-5\cdot0$	$-8\cdot0$	$+21\cdot6$	$+15\cdot2$	$+5\cdot6$
ii. $\frac{7}{16}$ $\frac{3}{8}$ 0	$+12\cdot6+i\,(0)$	$-0\cdot4$	$-8\cdot0$	$+13\cdot4$	$+5\cdot6$	$-2\cdot4$, $-5\cdot6$
iii. $\frac{3}{8}$ $\frac{3}{8}$ 0	$+15\cdot2+i\,(0)$	$+15\cdot2$	$-7\cdot2$	$+8\cdot0$	0	-8
iv. $\frac{5}{16}$ $\frac{3}{8}$ 0	$+5\cdot2+i\,(0)$	$+18\cdot2$	$-6\cdot0$	$+2\cdot4$	$-5\cdot6$	$-13\cdot7$, $+5\cdot7$
v. $\frac{5}{16}$ $\frac{5}{16}$ 0	$+0\cdot8+i\,(0)$	$+21\cdot6$	$-5\cdot0$	$-3\cdot2$	$-7\cdot2$	$-5\cdot7$
vi. $\frac{3}{8}$ $\frac{3}{8}$ $\frac{1}{16}$	$+10\cdot4+i\,(4\cdot0)$	$+8\cdot4$	$-6\cdot4$	$+5\cdot6$	0	$-5\cdot7$
vii. $\frac{5}{16}$ $\frac{3}{8}$ $\frac{1}{16}$	$+2\cdot8+i\,(7\cdot4)$	$+10\cdot2$	$-5\cdot8$	0	$-4\cdot0$	$-5\cdot7$, $+5\cdot7$
viii. $\frac{5}{16}$ $\frac{5}{16}$ $\frac{1}{16}$	$0+i\,(12\cdot4)$	$+12\cdot4$	$-4\cdot4$	$-5\cdot6$	$-4\cdot0$	0
ix. $\frac{5}{16}$ $\frac{5}{16}$ $\frac{1}{8}$	$-0\cdot8+i\,(10\cdot4)$	$-4\cdot4$	$-3\cdot2$	$-11\cdot2$	$+4\cdot0$	$+5\cdot7$
x. $\frac{5}{16}$ $\frac{1}{4}$ $\frac{1}{16}$	$-2\cdot8+i\,(10\cdot4)$	$+2\cdot8$	$-2\cdot8$	-8	$-4\cdot0$	$+5\cdot7$, $-5\cdot7$
xi. $\frac{5}{16}$ $\frac{3}{16}$ $\frac{1}{16}$	$0+i\,(6\cdot8)$	$-6\cdot8$	$-1\cdot2$	$-5\cdot6$	$-4\cdot0$	0
xii. $\frac{5}{16}$ $\frac{1}{8}$ $\frac{1}{16}$	$+2\cdot8+i\,(7\cdot4)$	$-4\cdot6$	$+0\cdot2$	0	$-4\cdot0$	$-5\cdot7$, $+5\cdot7$
xiii. $\frac{5}{16}$ $\frac{1}{16}$ $\frac{1}{16}$	$0+i\,(6\cdot8)$	$+6\cdot8$	$+1\cdot2$	$+5\cdot6$	$-4\cdot0$	0
xiv. $\frac{1}{4}$ $\frac{1}{16}$ $\frac{1}{16}$	$+6\cdot8+i\,(6\cdot0)$	$+1\cdot2$	$+6\cdot0$	$+3\cdot2$	$-3\cdot2$	$+5\cdot7$
xv. $\frac{3}{8}$ $\frac{1}{16}$ $\frac{1}{16}$	$-6\cdot8+i\,(4\cdot2)$	$+5\cdot4$	$-4\cdot4$	$+11\cdot2$	$+4\cdot0$	$-5\cdot7$

* The value for (411) contains, in general, both cosine and sine terms ($i = \sqrt{-1}$).

Arrangement 16.—In this case the atoms are so placed that a very large reflexion from (330) and (411) is an impossibility, so that no further consideration is required.

Arrangement 17.—The contribution of the 34 A. B. C and X atoms to the S.A. of (200) is positive for all values of a, b, c. Hence the contribution of the D atoms to the structure amplitude of (200) should be negative. Table V shows that only five sets of parameter values answer this condition, namely, (V), (VIII), (IX), (X) or (XI). In each of these five cases, a large value of (330) and (411) combined with small values for the planes (200), (220), (420) and (440) is obtained only for $a = \frac{1}{3}$ approx., $c = \frac{1}{2}$ approx. (420) and (440) also restrict b to the neighbourhood of $\frac{1}{3}$, so that all the parameter values are approximately fixed. However, with D atoms having the values (V), (VIII) or (IX), the structure amplitudes of (110) and (310) are quite large. The only way to get small values for these reflexions is to put $b = \frac{1}{4}$, which would give large values for (420) and (440). Values (X) and (XI) both give fairly satisfactory agreements for (110) and (310), but (211) is much too large in proportion to the values of (330) and (411). Table VI shows figures for the best set of parameter values, but clearly there is no sign of any correspondence in the intensity values.

Table VI.—NS^2 $(h^2 - k^2 - l^2)$ for Arrangement 17, with the best possible Parameter Values.

hkl.	Parameter Values of D atoms.					Observed Reflexion.	
	V.	VIII.	IX.	X.	XI.	Powder Photogram.	Precision Photogram.
110	15	7	4	3	1	absent	—
200	4	2	0	1	2	absent	—
211	—	—	—	6	15	absent	weak.
220	1	0	12	0	0	absent	weak.
310	6	7	9	2·5	0	absent	very weak.
330, 411	70	50	15	24	10	very strong	very strong.
420	0	1	7	4	1	absent	absent.
332	—	—	—	—	1	strong	strong.
440	1	1	1	—	1	absent	absent.

Arrangement 18.—Similar reasons limit the possible parameter values to (V), (VIII), (IX), (X) or (XI). Again the values of a and c must be approximately $\frac{1}{3}$. This would mean two sets of c atoms with the same co-ordinates, and is therefore quite impossible. Varying the values of c would produce large reflexions where nothing is visible.

Arrangement 19.—The parameter values $b = \frac{1}{3}$, $a = \frac{1}{3}$ give the smallest reflexions for (220), (420) simultaneously, but in each case the contribution of

the A, B and Y atoms is strongly positive. Consequently only those values of the D parameters which give negative contributions in both cases are possible. Only case (V) answers this requirement. However, this is not consistent with a small reflexion from (310). In short no set of parameter values gives a fit for all three reflexions simultaneously.

Arrangement 20.—As with arrangement 19, and for the same reasons, only case (V) is possible. A certain measure of agreement is obtained for $a = \frac{1}{3}$, $c = \frac{1}{3}$. However, (211) is far too large whilst (332) is much too small. These parameter values are therefore far from satisfactory. It is however not possible to get a better agreement. Table VII shows the best agreement which is possible for this arrangement. Clearly, it is still a long way removed from the correct structure.

Table VII.—$NS^2/(h^2+k^2+l^2)$ for Arrangement 20, D atoms having Parameter Values (v).

$h\,k\,l$.	$NS^2/(h^2+k^2+l^2)$.	Observed Reflexion.	
		Powder Photogram.	Precision Photogram.
110	4·7	absent	absent.
200	0	absent	absent.
211	6·2	absent	weak.
220	0·8	absent	weak.
310	1·9	absent	very weak.
222	0·4	absent	weak.
411 } 330 }	115	very strong	very strong.
420	2·8	absent	absent.
332	0·9	strong	strong.
440	2·3	absent	absent.

The above analysis shows that none of the arrangements 1–20 can satisfactorily account for the observed intensities. We are therefore forced to the conclusion that the structure of manganese must correspond to arrangement 21.

7. *Determination of the Parameters.*

The analysis of section 6 has left only one possible arrangement for the manganese atoms. The determination of the parameter values, of which there are 7, is a somewhat lengthy process but may be summarised briefly as follows. The contribution of the various atoms must be adjusted to give extremely small reflexions for a number of planes, whilst giving a very large reflexion to

(330) and (411). The parameter values which satisfy this condition are, following the notation of Table V, limited to the following possibilities. Of the two sets of D atoms, one must have parameter value III, V, VI, VII or VIII and the other values XIII, XIV, or XV. The value of the A parameter is necessarily about $\frac{1}{3}$. The possible values of the D parameters are all very much alike. The mean of all the above values for the D_1 atoms is $x_1 = 0{\cdot}34$, $y_1 = 0{\cdot}35$, $z_1 = 0{\cdot}04$ and for the D_2 atoms $x_2 = 0{\cdot}31$, $y_2 = 0{\cdot}06$, $z_2 = 0{\cdot}06$. A more exact evaluation of these parameters was attempted by trying the effect of a small variation in each of them in succession and so finding which combination of parameters gave the best agreement with observation. It appears that the most probable values of the D parameters are $x_1 = 0{\cdot}356$, $y_1 = 0{\cdot}356$, $z_1 = 0{\cdot}042$, $x_2 = 0{\cdot}278$, $y_2 = 0{\cdot}189$, $z_2 = 0{\cdot}189$. The value of the A parameter is $a = 0{\cdot}317$.

The space group is either T^3 or T^3_d. No other space group could have atoms distributed even approximately in this way, and could not give agreement between the observed and calculated intensities. The only difference between the atomic positions in T^3 and T^3_d is that the latter group requires two of the three parameter values of each D group to be equal. As far as we can tell, this is actually the case, and indicates that the space group is T^3_d, the crystal being therefore hemihedral. If there is any deformation which would lower the symmetry, it is so small that it scarcely affects the intensity values. It is unlikely that such a small effect would be revealed even if single crystals were available for the purpose of taking a Laue photograph.

Description of the Structure.

The structure of α-manganese has 58 atoms per unit cell with the following co-ordinates :—

2 X atoms $(0\,0\,0)$ $(\frac{1}{2}\,\frac{1}{2}\,\frac{1}{2})$

8 A atoms $(a\,a\,a)$ $(a\,\bar{a}\,\bar{a})$ $(\bar{a}\,a\,\bar{a})$ $(\bar{a}\,\bar{a}\,a)$ $(\frac{1}{2}+a, \frac{1}{2}+a, \frac{1}{2}+a)$ $(\frac{1}{2}+a, \frac{1}{2}-a, \frac{1}{2}-a)$ $(\frac{1}{2}-a, \frac{1}{2}+a, \frac{1}{2}-a)$ $(\frac{1}{2}-a, \frac{1}{2}-a, \frac{1}{2}+a)$

24 D_1 atoms $(d\,d\,e)$ $(d\,\bar{d}\,\bar{e})$ $(\bar{d}\,d\,\bar{e})$ $(\bar{d}\,\bar{d}\,e)$
$(\frac{1}{2}+d, \frac{1}{2}+d, \frac{1}{2}+e)$ $(\frac{1}{2}+d, \frac{1}{2}-d, \frac{1}{2}-e)$ $(\frac{1}{2}-d, \frac{1}{2}+d, \frac{1}{2}-e)$ $(\frac{1}{2}-d, \frac{1}{2}-d, \frac{1}{2}+e)$
$(d\,e\,d)$ $(\bar{d}\,\bar{e}\,d)$ $(d\,\bar{e}\,\bar{d})$ $(\bar{d}\,e\,\bar{d})$
$(\frac{1}{2}+d, \frac{1}{2}+e, \frac{1}{2}+d)$ $(\frac{1}{2}-d, \frac{1}{2}-e, \frac{1}{2}+d)$ $(\frac{1}{2}+d, \frac{1}{2}-e, \frac{1}{2}-d)$ $(\frac{1}{2}-d, \frac{1}{2}+e, \frac{1}{2}-d)$
$(e\,d\,d)$ $(\bar{e}\,d\,\bar{d})$ $(\bar{e}\,\bar{d}\,d)$ $(e\,\bar{d}\,\bar{d})$
$(\frac{1}{2}+e, \frac{1}{2}+d, \frac{1}{2}+d)$ $(\frac{1}{2}-e, \frac{1}{2}+d, \frac{1}{2}-d)$ $(\frac{1}{2}-e, \frac{1}{2}-d, \frac{1}{2}+d)$ $(\frac{1}{2}+e, \frac{1}{2}-d, \frac{1}{2}-d)$.

24 D_2 atoms $(f\,g\,g)$ $(f\,\bar{g}\,\bar{g})$ $(\bar{f}\,g\,\bar{g})$ $(\bar{f}\,\bar{g}\,g)$
$(\frac{1}{2}+f, \frac{1}{2}+g, \frac{1}{2}+g)$ $(\frac{1}{2}+f, \frac{1}{2}-g, \frac{1}{2}-g)$ $(\frac{1}{2}-f, \frac{1}{2}+g, \frac{1}{2}-g)$
$(\frac{1}{2}-f, \frac{1}{2}-g, \frac{1}{2}+g)$
$(g\,g\,f)$ $(\bar{g}\,\bar{g}\,f)$ $(g\,\bar{g}\,\bar{f})$ $(\bar{g}\,g\,\bar{f})$
$(\frac{1}{2}+g, \frac{1}{2}+g, \frac{1}{2}+f)$ $(\frac{1}{2}-g, \frac{1}{2}-g, \frac{1}{2}+f)$ $(\frac{1}{2}+g, \frac{1}{2}-g, \frac{1}{2}-f)$
$(\frac{1}{2}-g, \frac{1}{2}+g, \frac{1}{2}-f)$
$(g\,f\,g)$ $(\bar{g}\,f\,\bar{g})$ $(\bar{g}\,\bar{f}\,g)$ $(g\,\bar{f}\,\bar{g})$
$(\frac{1}{2}+g, \frac{1}{2}+f, \frac{1}{2}+g)$ $(\frac{1}{2}-g, \frac{1}{2}+f, \frac{1}{2}-g)$ $(\frac{1}{2}-g, \frac{1}{2}-f, \frac{1}{2}+g)$
$(\frac{1}{2}+g, \frac{1}{2}-f, \frac{1}{2}-g)$.

The values of the five parameters are—

$a = 0{\cdot}31_7$. $f = 0{\cdot}08_9$.
$d = 0{\cdot}35_6$. $g = 0{\cdot}27_8$.
$e = 0{\cdot}04_2$.

The space group is apparently T^3_d.

The structure of manganese is depicted in fig. 1. The simplest conception of the structure is obtained by regarding it in the following way. The basis of

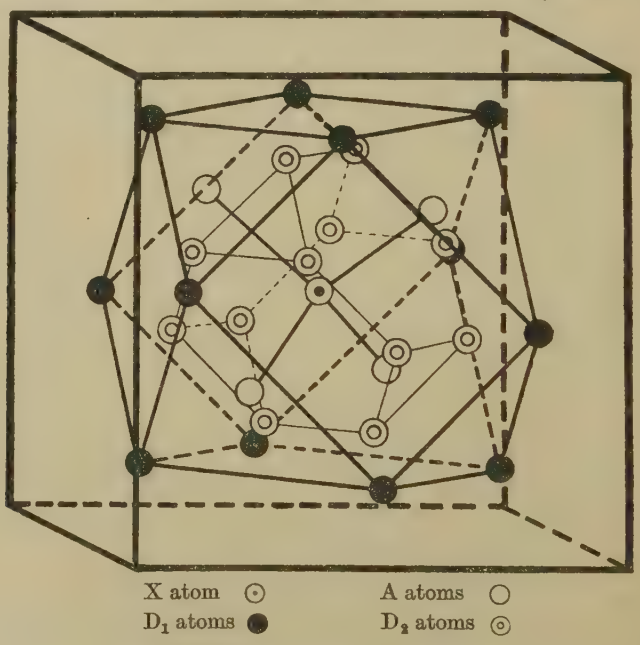


Fig. 1.—The structure of α-Manganese.

the whole arrangement is one single body-centred cubic lattice, but each lattice point is represented by a cluster of 29 atoms. One such cluster is shown in the figure, surrounding the central atom of the unit cube. The structure being

body centred there are corresponding clusters of atoms around the corners of the cube, but it would be too confusing to attempt to insert these in the figure. One point must be emphasised. These clusters of atoms are in no way chemical molecules or even groups. An atom is no more related to the other atoms within the cluster than to neighbouring atoms outside the cluster. The cluster is in fact a mere geometrical conception, serving as an aid to the imagination.

Atoms of four different types are shown in the drawing. At the centre lies a single atom X. This is surrounded by a framework of atoms D_2. These atoms are arranged in such a way that they build up an octahedron in which opposite faces are of different sizes so that the symmetry is tetrahedral. A little further out are four A atoms arranged tetrahedrally around the centre. Lastly, the twelve D_1 atoms build up a polyhedron with cubic and octahedral faces, the latter being of two different sizes. The whole cluster of atoms has therefore tetrahedral symmetry, which is also the symmetry of the crystal taken as a whole.

The structure contains atoms of four different kinds. In each unit cell there are 2 X atoms, 8 A atoms, 24 D_1 atoms and 24 D_2 atoms, making 58 in all.

Table VIII.—Interatomic Distances of α-Manganese.

D_2 atoms.			D_1 atoms.		
No. of Neighbours.	Type of Neighbours.	Distance.	No. of Neighbours.	Type of Neighbours.	Distance.
1	D_2	2·24	1	D_2	2·45
2	D_2	2·38	1	A	2·49
1	D_1	2·45	2	D_2	2·51
2	D_1	2·51	2	D_2	2·66
2	D_1	2·66	6	D_1	2·67
2	A	2·69	1	A	2·96
1	X	2·71			
1	A	2·89			
X atoms.			B atoms.		
No. of Neighbours.	Type of Neighbours.	Distance.	No. of Neighbours.	Type of Neighbours.	Distance.
12	D_2	2·71	3	D_1	2·49
4	A	2·82	3	D_2	2·69
			1	X	2·82
			3	D_2	2·89
			3	D_1	2·96

The atoms of the various types differ with respect to the relative position of their neighbouring atoms as Table VIII above shows.

Figs. 2 and 3 represent sections of the unit cell drawn through the centre parallel to (100) and ($\bar{1}10$) respectively. In order to give a concrete picture

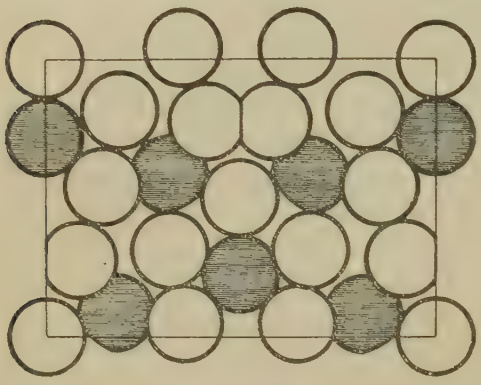


Fig. 2.—A Section through the Structure of α-Manganese parallel to (110).

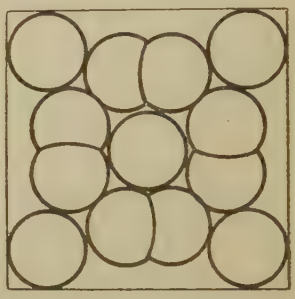


Fig. 3.—A Section through the Structure of α-Manganese parallel to (100).

of the structure the atoms are represented as spheres of radius 2·5 Å. This is of course a very approximate representation of the atomic packing as many of the interatomic distances vary considerably from this amount, being different in different directions around each atom. This fact is shown particularly by the D_2 atoms which must be represented as being slightly truncated in order to fit them together. They occupy decidedly less room than the other atoms in the structure. The X atoms on the other hand occupy a larger volume. This is clearly shown in fig. 2, the central atom being separated from its nearest neighbours. These differences in atomic dimensions must be closely related to the complex nature of the structure. They may perhaps be explained by supposing that the electrons are not shared equally between the different atoms, so that some have a surplus and some a deficit of electrons. This implies the existence of something more akin to a compound than to a true element in the α-manganese structure. The best analogy to such a structure would be found in the alloys, the structure of γ-brass*† (Cu_5Zn_8) in particular being strikingly like that of manganese.

* A. Westgren and G. Phragmén, 'Phil. Mag.,' vol. 50, p. 311 (1925).

† A. J. Bradley and J. Thewlis, 'Roy. Soc. Proc.,' A, vol. 112, p. 678 (1926).

Summary.

(1) The structure of α-manganese has been deduced from the data of Westgren and Phragmén, who had shown that it was cubic, the lattice dimensions being 8·894 A.

(2) α-Manganese contains 58 atoms per unit cell.

(3) The space group is T^3_d. There are four sets of equivalent positions, containing respectively 2, 8, 24 and 24 atoms.

(4) The exact position of the atoms is defined by five parameters, which have been evaluated.

(5) The structure is based on a single body centred cubic lattice but each lattice point is replaced by a cluster of atoms, with tetrahedral symmetry.

(6) The interatomic distances range from 2·25 Å to 2·95 Å, indicating an unequal distribution of electrons between the various atoms.

The authors express their thanks to Prof. W. L. Bragg, F.R.S., for his kind interest and help in this investigation ; to Dr. A. Westgren and Mr. G. Phragmén for their kindness in making available their data ; to Mr. A. P. M. Fleming, O.B.E., M.Sc. (Tech.), Director of Research of Metropolitan Vickers Electrical Co., Ltd., for permission to publish the results ; and to the Royal Commissioners of the Exhibition of 1851 for a scholarship which enabled the work to be completed. Thanks are also due to the Board of Scientific and Industrial Research for a grant to one of us.

The Rusting of Steel Surfaces in Contact.

By G. A. Tomlinson.

(Communicated by Sir William Hardy, F.R.S.—Received May 28, 1927.)

When two hard steel surfaces are subject to relative motion under some load, it is well known that after some time a brown stain is liable to form at the common surface. A common example occurs in the case of a micrometer anvil operating on a spherical abutment, where it is periodically necessary to clean the surfaces in contact, owing to a considerable reddish brown deposit having accumulated.

In engineering work the same phenomenon is found but in a more acute form. When two machined steel surfaces are held firmly in contact and at the same time are subject to vibration, it is often found on taking them apart that the surfaces have become cemented together by the production of relatively large quantities of oxide, and the individual surfaces are badly pitted and have a corroded appearance. This action goes on, it should be noted, without relative motion of the surfaces, or more correctly, without relative motion of ordinarily measurable amounts, a question which is taken up later.

Oiling of the surfaces before assembly certainly does not prevent this effect, and there is little direct evidence as to whether it is beneficial to a limited extent or not at all.

This effect becomes of importance in any machine which has to be frequently re-assembled as it may be found that highly finished ground surfaces of certain parts are seriously deteriorated by the action described.

The phenomenon was mentioned by Eden, Rose and Cunningham* who suggested that it was associated with the application of alternating stress. After this, in an editorial article in the 'Engineer,'† followed by a discussion in the correspondence columns, the subject was taken up more fully. Various explanations of the rusting were suggested including the adsorption of moisture, electrolysis and intense surface heating. Since this time, although the effect has become of more frequent occurrence with the development of high-speed machines, the writer has found no further references to the possible cause.

In order to discover the cause of the rusting which occurs under the circumstances described, a series of experiments were made involving a plane surface,

* 'Proc. Inst. Mech. Eng.,' p. 875 (1911).

† 'Engineer' (Nov. 24th, 1911).

and a spherical surface in relative motion at their common point of contact, the motion in some cases being a rotation about the common normal and in others a tangential sliding motion.

In the rotation experiments a bar of wood was pivoted on the sphere which rested on a slip gauge. The bar carried a mass at each end below the level of the sphere, so that the centre of gravity was well below the point of support. The whole was set in rotation by a magnetic couple. After a few revolutions the plane surface was examined under a metallurgical microscope. A typical experiment is described below :—

$\frac{1}{16}$-inch Hoffmann ball on a lapped slip gauge, both of hardened steel.
Surfaces mechanically cleaned.
Load : 13 ozs. Mean intensity of pressure 223,000 lbs. per sq. in.
Number of revolutions about 8.

Afterwards a minute brownish patch could be seen with the naked eye, which under the microscope showed considerable oxidation debris having bright reddish and brown colours. In a number of places the debris has the appearance of having been spread into a film by the motion and pressure, these places having a bright blue or green tint surrounded by a dark brown border. The patch of oxidation is fairly permanent and cannot be rubbed off with a cloth. A photograph of this particular result is shown in fig. 1.

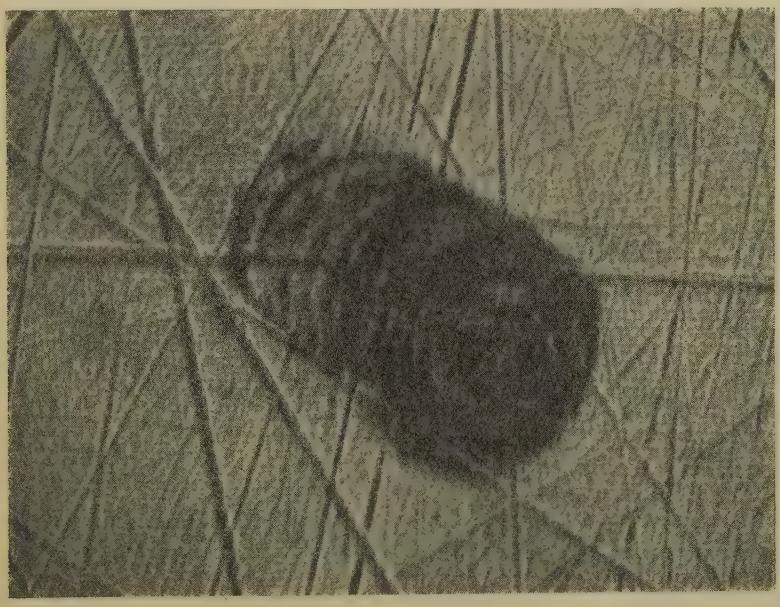


Fig. 1. (× 450)

A single passage of one surface over the other is sufficient to produce a track of oxidation debris that is plainly visible. This is in fact clear from fig. 1 in which the rotating ball has by chance progressed slightly. At one end of the elongated patch there is a rough spiral form showing single passages of the one surface over the other.

Two important factors, the effect of which has been examined in various experiments, are the pressure between the surfaces and the condition of the surfaces.

It was not possible with the rotating bar to obtain very low pressures and this method was therefore given up in favour of a tangential sliding motion of the sphere over the plane. Sometimes a rotating plate has been used with the sphere stationary at the end of a balanced lever to which any required load could be added. In this case the relative motion occurs in small circles. With another form of apparatus the plate is fixed and the sphere, again carried by a weighted lever on pivots, is moved to and fro through a short distance.

In the earlier experiments, the steel slip was lapped and the surfaces were merely mechanically cleaned. Later, only surfaces highly polished on wet chamois leather with chromium oxide were used, and when clean surfaces were required they were washed with an alcohol-ether-ammonia cleaning mixture.

From a considerable number of different experiments, a few significant cases have been selected and these are shown for convenience in tabular form in Table I.

In experiment A no rusting was produced and as the surfaces were only cleaned in a mechanical way by careful wiping, it is most probable that a persistent primary film remained and completely separated the metal surfaces. In the experiment cited above in which copious oxidation was produced the mean contact pressure was over seven times as high as in A, suggesting that in this case the stress was more than the film could support without rupture.

Experiment D is of considerable interest. A load of only 0·23 grammes on a comparatively large sphere was found to produce the typical oxidation debris to quite a marked extent when the surfaces were well washed with the cleaning fluid and the primary film was presumably removed. This experiment disposes of one possible explanation of the rusting which might be suggested by ordinary observation of the phenomenon as it occurs in practice when lapped surfaces are found to rust on being rubbed together. This explanation is that the surfaces, which on a minute scale are completely scored by the lapping abrasive, are crushed together under a heavy normal stress and a fine powder is thus produced by a mutual grinding action, the powder oxidising to form the well-

Table I.

	Nature of surfaces.	Diameter of sphere.	Load.	Mean intensity of pressure.	Motion applied.	Result.
		inches.	gms.	lbs. per sq. in.		
A	New Hoffmann ball on lapped slip. Cleaned mechanically	0·344	25	30,000	Rotation about common normal	No visible effect.
B	Hoffmann ball on soft steel plate. Cleaned mechanically	0·344	213	61,000	Do.	Surface much torn up, but no oxidation colour to be seen.
C	Hoffmann ball on lapped hard steel. Cleaned mechanically	0·0325	368	354,000	No motion. Load left on for 40 hours	No effect.
D	Hoffmann ball highly polished on hard steel highly polished. Surfaces cleaned chemically	0·469	0·23	5,000	To and fro about 100 times	Considerable brownish debris, partly spread on the surface and partly swept up at the ends of the track. The latter could be wiped away.
E	Two polished Hoffmann balls on polished hard steel. Cleaned chemically	0·14	180 per ball	106,000	Pure rolling to and fro. 8,000 complete oscillations	No effect.
F	Fused glass sphere on polished steel. Cleaned chemically	0·25	4·6	12,500	To and fro about 100 times	Considerable amount of oxidation residue of reddish colour. Cannot be removed by repeated wiping.
G	Polished stellite sphere on polished steel. Cleaned chemically	0·16	12·3	40,000 (approx.)	Tangential motion in small circles	Typical brown debris produced. Visible with the naked eye.
H	Polished agate sphere on polished steel. Cleaned chemically	0·16	2·3	—	To and fro about 150 times	Considerable oxidation produced.
J	Polished steel sphere on polished steel. Surface smeared with vaseline	0·14	12·3	43,500	Tangential motion in small circles	No effect.
K	Do. do.	0·14	110	90,000	Do.	Marked rings of brown oxidation produced which are permanent after wiping off the vaseline.
L	Polished steel sphere on polished steel. Surface smeared with castor oil	0·469	365	60,000	To and fro motion	Usual oxidation debris produced but not continuous along the whole track, mostly at one isolated place.
M	Do. do.	0·469	45·6	30,000	Do.	No oxidation visible.

known stain. In this experiment the surfaces are both highly polished and no scratches at all are visible under the microscope. The mean contact stress is only 5,000 lbs. per square inch whereas experiment C shows that similar surfaces can bear a normal stress as high as 354,000 lbs. per square inch without any injury visible under the microscope. It thus seems most improbable that in

this case the oxidation is the result of any such coarse abrasive action between the surfaces. Experiment F also strongly confirms this view, as here the spherical surface is a fused bead which is comparable in texture with a fluid surface and certainly will have no prominences such as are necessary for abrasion to take place.

Experiment C shows that no oxidation occurs unless there is relative tangential motion of the surfaces, even under excessive intensity of normal stress. Hence the effect is not due to the local application of heavy pressure which might conceivably set up an electrolytic action. This is further confirmed by experiment E in which a motion of pure rolling was used, again with a fairly heavy normal stress. This motion produces no oxidation.

Experiments F, G and H show that the same result follows from the sliding contact of glass, stellite and agate, in all cases with comparatively small contact stresses.

The most probable explanation of the effect which suggests itself is that it is a result of molecular cohesion. When two solids touch the forces between the molecules or at least some of the molecules are sufficiently high to cause the molecule to be detached by a lateral movement. It seems certain that the force between two molecules which approach and recede normally is definitely smaller than the forces holding either molecule to the solid. Experiment E in which the relative motion is always exactly normal to both surfaces shows this, and all experience with ball bearings proves that the rusting effect does not occur with pure rolling. The nature of the bonds between the molecules appears to be such that the cohesive force of a visiting molecule is quite insufficient to pluck the molecule out normally, but is sufficient to detach it from the solid when applied tangentially. This suggests the view that the boundary molecules, by virtue of their unsymmetrical position, have a considerable degree of orientation, so that an external tangential force is able to disturb the initial equilibrium so much that individual molecules as distinct from finite particles can be wrenched away. To use a crude analogy, a tooth is more easily uprooted by a side pull than by a normal pull. The molecules so detached combine very quickly with oxygen molecules from the atmosphere.

It is not easy to obtain any direct experimental evidence that molecular cohesion takes place, although strong molecular forces would be expected when the boundaries of two solids are brought together within the range of the molecular fields. That molecular cohesion occurs when two clean solids touch seems to be certain. Sir W. Hardy has expressed the opinion that the static friction between clean surfaces is due to cohesion.

The behaviour of glass is interesting, as glass appears to possess a greater faculty for cohering than any other material examined, possibly owing to the comparative ease with which clean surfaces can be obtained. One simple experiment which is easily carried out is convincing that cohesion at once occurs in an actual contact. If a piece of plate glass and a fused glass bead at the end of a rod are both carefully cleaned, and the rod is poised lightly in the fingers and allowed to stroke the plate, there is felt to be a sequence of snatches as the bead welds on and breaks away repeatedly. If afterwards the plate is examined with a lens it is seen to be scored with fine dotted lines wherever the bead has touched it, even with the lightest touch. In the same experiment with hard steel surfaces the tactile sensation of welding is less marked, but if the plane surface is examined under the microscope the path followed by the ball is shown by a thin track of reddish oxidation debris.

The cohesive force between glass surfaces can be seen in operation by another simple experiment. Two fine glass threads are drawn out which immediately after drawing are probably perfectly clean. If one thread is made to touch the other the cohesion force between them can be plainly observed by the deflection of the two threads. If one thread is fairly stiff it may even support the whole weight of the finer thread by cohesion.

Two different experiments have been made which lend some support to the above theory of the oxidation effect. The experiments already described prove that with a tangential motion either molecules or grosser particles are torn from the steel surface and are very rapidly oxidised. If actual molecules are detached it would be reasonable to expect this to occur with very minute tangential displacements. In the experiments to be described the characteristic rusting effect was found to occur after a very minute motion had been repeated a number of times. In the first experiment a spherical face was caused to make repeated contacts at the same place on a plane polished steel face. After a number of contacts the exact place was located by making a small grease spot at equal distances on each side, and the specimen was examined under the microscope.

The action produces the typical oxidation effect though to a much smaller extent than in the experiments in which there is considerable sliding. The effect, however, especially in view of the exact location of the place, is quite unmistakable. Steel, glass, stellite and agate all produce the result on steel.

In these experiments there is most probably a minute lateral displacement as the load comes on to the sphere, due to the elastic bending of the apparatus. This slip must be very small though unfortunately it is not computable, but

it is sufficient on the view that molecules are torn away individually by cohesion forces.

In the second experiment a known very small displacement was produced and repeated a large number of times. A light T-shaped frame of wood, A, was supported on three Hoffmann balls, B, C, and D (fig. 2) on slip gauges. The ball,

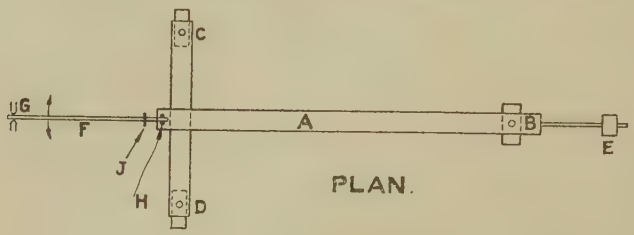


Fig. 2.

B, and the plane surface on which it bears were highly polished, and practically the whole of the weight is adjusted by the counter weight, E, to be on this ball. the other balls having only just enough load to keep them in contact. By means of the rod, F, pivoted on a vertical spring steel strip, J, the frame can be given small angular oscillations about the centre of contact of ball B. The rod, F, oscillated between stops, G, at one end while the other end is placed between two pins, H, on the frame. These pins are at a distance of 10 inches from the ball B, and the amplitude of the angular oscillation can be found from the distance between the stops, G. The radius of the circle of contact can be calculated and hence the maximum linear displacement of the surfaces can be found.

In a particular experiment the following values were used :—

Load on the ball	113 grammes
Diameter of ball	0·03125 inches
Radius of circle of contact	0·00055 inches
Displacement of frame at the pins	0·0016 inches
Total number of oscillations	50,000 (approx.)

Hence the displacement of the two surfaces at a radius 0·00055 inches is

$$0{\cdot}0016 \times \frac{0{\cdot}00055}{10} = 8{\cdot}5 \times 10^{-8} \text{ inches.}$$

In the above experiment a small but definite amount of oxidation debris was afterwards observed at the place of contact. A large number of these minute oscillations must be made to produce a clearly visible effect. In this case, for example, about 21,000 oscillations give an aggregate motion equal to one revolution. Further, when the surfaces have a slight oscillatory motion, there

will be a tendency for them to become to some extent separated by the oxide which has not the same chance of being swept off and left behind as when the motion is continuous.

Two further experiments were made in which the amplitudes of the motion were $6{\cdot}5 \times 10^{-8}$ inches and $3{\cdot}5 \times 10^{-8}$ inches. In the latter case a thorough search showed no visible trace of oxidation. With the intermediate amplitude, after some searching, a slight surface discolouration was detected, so slight compared with the result obtained with the larger amplitude, that the opinion formed that it was oxidation may be wrong. We may, therefore, say that with a maximum relative motion of $6{\cdot}5 \times 10^{-8}$ inches, the detaching of molecules is practically arrested. There is thus experimental evidence that there is a certain magnitude of relative motion below which the molecules may be disturbed and yet return to their original configuration. With a greater motion some, at least, are not able to recover their stable position.

It is a point of great interest that the value of this critical displacement is near the value of the radius of molecular attraction, found to be 8×10^{-8} inches by Johonott and by Chamberlain independently, from observation of the minimum thickness of a liquid film. The radius of attraction is of the same order for molecules of widely different molecular weights, and there is no reason to suppose that the molecular field is materially different in the liquid and solid states.

These results also make it clear how copious rusting may occur at certain contact surfaces in high-speed machines, although apparently the surfaces are secured rigidly together. Actually there is most probably a minute to and fro motion occurring as the rust is invariably found associated with vibration. With high frequency vibration the aggregate motion per hour may be considerable, hence the large accumulation of oxide found after a length of time.

The experiments which have been described not only support the theory that the rusting is a consequence of molecular cohesion, but they also afford evidence which is contrary to certain other suggested explanations.

If we consider, first, electrolysis in the stressed material, this should occur irrespective of any motion of the surfaces, whereas tangential motion was found to be necessary. Further, electrolysis requires an appreciable time, and it was found that a single quick passage of one surface over the other produced rust. Similar arguments apply against the view that the rusting is a result of an adsorbed film of moisture.

The possibility of local heating to a high temperature needs a little more consideration. Cases may no doubt occur in which the heat generated can only

be diffused by conduction into the body of the metal by a large temperature gradient at the contact surface. Polishing affords a somewhat analogous case, and it is considered by some that polishing is due to surface fusion.* If, however, we consider the experiment in which rusting was produced by small oscillations, it appears to be impossible that any very appreciable temperature rise could have occurred. The rate of loss of energy can be calculated fairly accurately, and, in this experiment, it amounts to about 542 ergs in a period of 330 minutes from which the rate of generation of heat is $6 \cdot 5 \times 10^{-10}$ calories per second.

The radius of the circle of contact is 0·00055 inch, and assuming that all the heat generated must pass from this contact area by conduction, the necessary temperature gradient normal to the contact surface is computed to be approximately 0·001° C. per centimetre. Hence, only a very minute rise of temperature can be possible, and it may be concluded that the rusting effect is not dependent on the production of a high surface temperature.

In the last four experiments given in Table I the surfaces were covered with a smear of vaseline or castor oil. In each case oxidation did not occur with comparatively low normal stress, and did occur when the stresses were higher. A film of great mechanical strength separated the surfaces and prevented molecular approach of the solids, but there is a limit to the strength of such a film, and a sufficiently high normal stress at the contact of the sphere and plane was able to rupture the film. This is of practical importance as in actual apparatus the film is most probably present, and, therefore, the conditions governing the stability of the film were studied in more detail.

For this purpose glass surfaces appeared to be the most suitable for several reasons. In the first place to use the oxidation effect on steel would be a very slow method. The work of Sir W. B. Hardy has shown that the static friction of the contact would afford a much more convenient indication of the condition of the film, and this method was employed. The coefficient of friction between a fused glass sphere and a glass plate flooded with castor oil was measured over a considerable range of normal stress. Hardy has given the coefficient of friction of clean dry glass as 0·94, and when lubricated with castor oil as 0·1. Glass is thus specially suitable, as rupture of the film permitting a partial cohesion of the solids would be expected to be followed by a decided increase in the coefficient. Further, with glass the ideal smooth spherical surface can be more nearly attained by fusion.

When steel is used there is reason to believe that the minute prominences left

* 'Nature,' September 4, 1926, *et seq.*

even after polishing may give rise to local stresses higher than the theoretical maximum as calculated from the Hertz equation. Certain results obtained using steel suggest that this is the case.

The results found with glass surfaces and castor oil are shown plotted in fig. 3, every point being a single observation. The coefficient of friction has been

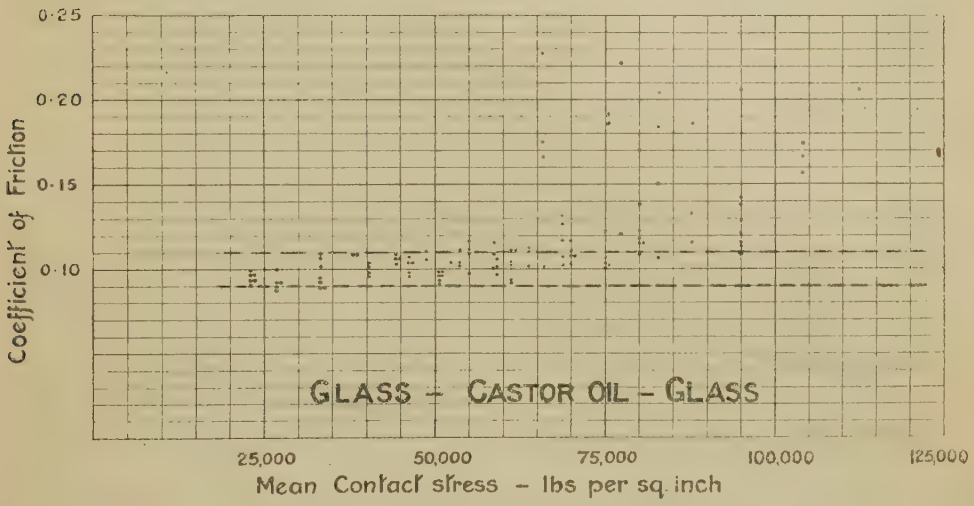


Fig. 3.

plotted with the mean normal stress found by calculating the radius of the circle of contact. In a few cases two observations coincide, and in order to show both on the diagram they have been placed plotted thus ·–· the left hand point being placed at the correct stress.

The first point of interest is that there is a critical stress intensity of about 65,000 lbs. per square inch at which the film appears to suffer rupture. Below this stress nearly all the observed values of the coefficient fall between limits 0·09 and 0·11, the mean being closely 0·10. With higher stresses a clear majority of the points lie well above 0·11. A few points, from 65,000 to 75,000, still lie between 0·10 and 0·11, but above this the coefficient is clearly much higher. Above the critical stress the values obtained are far more erratic. Thus, for example, eight observations at a mean stress of 95,000 cover a range from 0·109 to 0·205. The inference is that the rupture of the film is in irregular patches, the form and size of which vary widely in different experiments. The film probably does not vanish completely over a central circular area where the intensity of pressure exceeds a critical value, as the thinnest films exhibit the characteristics of a solid rather than of a fluid. Most probably there are various

places at which the distance between the solids is slightly smaller than the average distance, and it is at these points that the rupture commences. It is thus not necessary to suppose that a certain load which involves a mean stress of 95,000 is supported in any definite proportion by the molecules of the fluid and of the solid. The fluid film, where the solid boundary surfaces are very parallel, is probably capable of withstanding much greater stresses than the mean stress over the area. Hence the actual observed value 0·109 of the coefficient in one case with a mean stress of 95,000. It should be mentioned that a fresh portion of both the surfaces was used in every experiment, as the surface is injured locally when cohesion occurs between the solids.

It would not be safe to assume that the film offers the same resistance to rupture when bounded by hard steel surfaces. The limiting film is in a peculiar physical condition as shown by its ability to withstand high normal stress, and there are likely to be strong molecular interactions between the solid and the fluid. Hence a different behaviour with steel surfaces would not be surprising. On account of the greater practical importance in the case of steel, a similar series of measurements was made using polished hard steel surfaces. The results are shown in fig. 4, and indicate that the rupturing stress is actually much

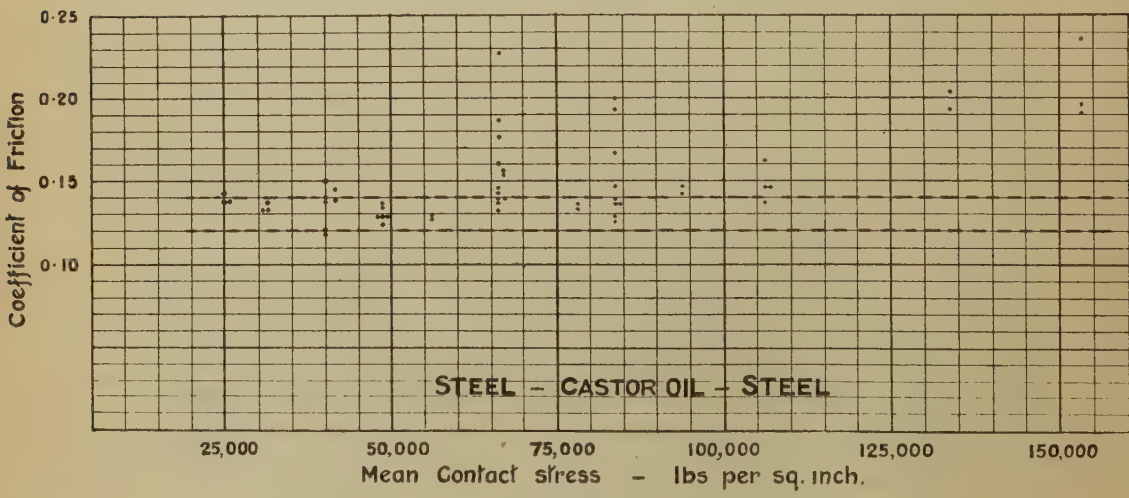


Fig. 4.

the same as with glass. The results found with lower stresses than the critical value are rather more erratic, and it is thought that occasionally a certain degree of solid cohesion occurs with stresses well below 65,000 lbs. per square inch for the reasons stated earlier in this section. In evidence of this, some experiments may be mentioned in which a smear of castor oil was used and the rupture

of the film was judged by a microscopic examination of the surface afterwards. It was found in certain cases with a stress as low as 30,000 that there had occurred some definite solid cohesion.

The writer wishes to thank Sir William Hardy for the kind interest he has taken in the work described, and Sir Joseph Petavel and Mr. J. E. Sears for their encouragement and facilities for carrying out the experiments at the National Physical Laboratory.

A Note on the Specific Heat of the Hydrogen Molecule.

By David M. Dennison, Ph.D., University of Michigan.

(Communicated by R. H. Fowler, F.R.S.—Received June 3, 1927.)

In a recent article F. Hund* has treated the problem of the specific heat of the hydrogen molecule on the basis of the wave mechanics. The total number of rotational states are divided due to the homopolar character of the molecule into two groups, to the one of which belong wave functions symmetrical in the two nuclei, and to the other wave functions which are antisymmetrical in the nuclei. Hund has suggested that the presence of both groups in hydrogen may be accounted for by assuming that the nuclei possess a spin, in which case transitions between symmetrical or between antisymmetrical states will have their usual intensity but transitions between symmetrical and antisymmetrical states will be very weak, of the order of the coupling of the nuclear spins. He then writes the following expression for the rotational specific heat,

$$\frac{C_r}{R} = \sigma^2 \frac{d^2}{d\sigma^2} \log Q,$$

$$Q = \beta [1 + 5e^{-6\sigma} + 9e^{-20\sigma} + \ldots] + 3e^{-2\sigma} + 7e^{-12\sigma} + 11e^{-30\sigma} + \ldots, \quad (1)$$

where $\sigma = h^2/8\pi^2 IkT$ and β is the ratio of the weights of the symmetrical group of states to the antisymmetrical group. Hund has found that he obtains a close agreement between (1) and the observed specific heat curve only when β has about the value 2, that is when the symmetrical states have twice the weight of the antisymmetrical. He further obtains for this case $I = 1{\cdot}54 \times 10^{-41}$ gm. cm.^{2}, the moment of inertia of the H_2 molecule.

These values for β and I are not in agreement with the observed features of the band spectra of H_2. A careful analysis of the far ultra-violet bands of H_2

* F. Hund, 'Z. f. Physik,' vol. 42, p. 93 (1927).

has been made by T. Hori* to whom I am greatly indebted for allowing me to see the manuscript of his work. Hori finds that the moment of inertia in the normal state has the value $I = 4{\cdot}67 \times 10^{-41}$, and that the transitions between antisymmetrical terms are about three times as strong as the corresponding transitions between symmetrical terms, that is $\beta \simeq \frac{1}{3}$. He does not find any lines corresponding to transitions between symmetrical and antisymmetrical terms. These values for β and I when set into (1) lead to a specific heat curve having a sharp and high maximum, in no way agreeing with the observed curve.

It now suggests itself that the difficulties encountered in comparing these sets of data may lie in the assumption that the symmetrical and antisymmetrical terms can combine, an assumption which determined the form of (1). The coupling of the nuclear spins with the spin of the molecule, which determines these transitions, will indeed be very small, much smaller than the coupling forces between the electronic spins and the orbits which give rise to the very weak transitions between ortho- and para-helium. Let us make the assumption that the time of transition between a state symmetrical in the rotation, and an antisymmetrical state is very long compared with the time in which the observations of the specific heat are made. In this case we have in effect two distinct gases, the one formed from the symmetrical states with a specific heat C_s and the other containing the antisymmetrical states with a specific heat C_a, where

$$\left.\begin{aligned} \frac{C_s}{R} &= \sigma^2 \frac{d^2}{d\sigma^2} \log Q_s, \quad Q_s = 1 + 5e^{-6\sigma} + 9e^{-20\sigma} + \ldots \\ \frac{C_a}{R} &= \sigma^2 \frac{d^2}{d\sigma^2} \log Q_a, \quad Q_a = 3e^{-2\sigma} + 7e^{-12\sigma} + 11e^{-30\sigma} + \ldots \end{aligned}\right\}. \tag{2}$$

The final rotational specific heat of the mixture is

$$\frac{C_r}{R} = \frac{\rho C_s + C_a}{(1+\rho)R}, \tag{3}$$

where ρ is the proportion of symmetrical to antisymmetrical molecules. In Table I there is given a series of values of C_s/R and C_a/R computed for the argument σ.

It is evident that since C_s/R rises to a maximum of about 1·5 and C_a/R rises steadily to unity that ρ must have a value less than 1 if the curve C_r/R is to fit the experimental curve. The following method was used to determine ρ. The observed values for the specific heat of H_2 as given by Eucken,† Scheel and

* T. Hori, 'Z. f. Physik' (*in print*).

† A. Eucken, 'Preuss. Akad. d. Wiss.,' p. 141 (1912).

Table I.

σ.	$\frac{C_s}{R}$	$\frac{C_a}{R}$	$\frac{\frac{1}{3}C_s + C_a}{\frac{4}{3}R}$	$T_{exp.}$	$\sigma T_{exp.}$	$T_{calc.}$
1·2	0·193	0·0021	0·050	70	84·0	71·2
1·0	0·436	0·0106	0·117	85	85·0	85·5
0·8	0·875	0·0501	0·256	104	83·2	106·9
0·7	1·142	0·1038	0·363	118	82·6	122
0·6	1·374	0·206	0·498	140	84·0	142
0·5	1·464	0·381	0·652	176	88·0	171
0·4	1·341	0·635	0·811	223	89·2	214
0·35	1·224	0·767	0·881	251	87·9	244
0·3	1·111	0·886	0·942	286	85·8	285

Heuse,* Giacomini,† Brinkworth‡ and Partington and Howe§ were plotted against the temperature. An averaging curve was drawn through these points and this curve was assumed to give the observed variation of the specific heat with temperature. A value of ρ was selected and C_r/R computed from (3) and for every point the value of T exp. was read off the observed curve. A column $\sigma T_{exp.}$ was then constructed which should consist of constant values if the computed curve agreed exactly with the observed curve. It was found that the best agreement occurred when $\rho = \frac{1}{3}$, that is when the antisymmetrical molecules are three times as numerous as the symmetrical ones. The values for C_r/R for this case are given in the fourth column of the table, followed by a column for $T_{exp.}$ and for $\sigma T_{exp.}$ (Although it is not possible to fix the value of ρ exactly from the present data, it can be shown that $\rho = \frac{1}{3}$ fits the observed curve distinctly better than $\rho = 1/3{\cdot}5$ or $\rho = 1/2{\cdot}7$.)

Taking the average of $\sigma T_{exp.}$ to be 85·5, we may find the computed values for the temperature as given in the seventh column of the table. The agreement between $T_{calc.}$ and $T_{exp.}$ is very satisfactory, the greatest deviation (less than 4 per cent.) is within the limit of the experimental error. If $\sigma T = 85{\cdot}5$ we find the moment of inertia of the hydrogen molecule to be $I = 4{\cdot}64 \times 10^{-41}$ gm. cm.^{2} in substantial agreement with the moment of inertia found by Hori from band spectra data.

In conclusion we may say that by assuming that the symmetrical and antisymmetrical rotational states of the hydrogen molecule do not combine during

* K. Scheel and H. Heuse, 'Ann. d. Phys.,' vol. 40, p. 473 (1913).

† F. A. Giacomini, 'Phil. Mag.,' vol. 50, p. 146 (1925).

‡ J. H. Brinkworth, 'Roy. Soc. Proc.,' A, vol. 107, p. 510 (1925).

§ J. R. Partington and A. B. Howe, 'Roy. Soc. Proc., A, vol. 109, p. 286 (1925).

a time long compared with the time of the experiment, we obtain a specific heat curve which follows the observed curve to within the errors of observation, and that moreover the constants ρ and I are in good agreement with the values of these constants as found in the band spectrum of H_2.

[*Added June* 16, 1927.—It may be pointed out that the ratio of 3 to 1 of the antisymmetrical and symmetrical modifications of hydrogen, as regards the rotation of the molecule, is just what is to be expected from a consideration of the equilibrium at ordinary temperatures if the nuclear spin is taken equal to that of the electron, and only the complete antisymmetrical solution of the Schrödinger wave equation allowed.*

While it would not appear possible to produce only the one modification of the molecule by a combination of two hydrogen atoms since the heat of dissociation of H_2 is so much higher than the difference between the first rotational states, other experiments might be performed which would show the non-combining character of these two sets of rotational states and possibly even allow them to be separated. Indeed, the far ultra-violet absorption spectrum of H_2 at low temperatures would show at once whether the molecules all go into the zero state of rotation or whether they remain in the zero and first rotational states in the ratio of 1 to 3 as is suggested in the present note.]

I wish to express my thanks to Mr. R. H. Fowler for much helpful criticism and to Prof. T. Hori for the opportunity of seeing the results of his work before their publication. I wish also to acknowledge with gratitude a stipend from the University of Michigan.

* W. Heisenberg, ' Z. f. Physik,' vol. 41, p. 239 (1927), in particular see p. 264.

Bakerian Lecture.—*A New Mass-Spectrograph and the Whole Number Rule.*

By F. W. Aston, F.R.S.

(Received June 16, 1927.)

[Plate 8.]

Introduction.

The original mass-spectrograph was set up in the Cavendish Laboratory in 1919. Its resolving power was sufficient to separate mass lines differing by about 1 in 130 and its accuracy of measurement was about 1 in 1000. These capabilities sufficed to determine with fair certainty the isotopic constitution of over 50 elements, and to demonstrate that, with the exception of hydrogen, the masses of all atoms could be expressed as integers on the scale O = 16 to one or two parts in one thousand. An account of these researches has already been published.* The instrument itself was not actually dismantled until March, 1925, but some years before then it had been realised that for advance in two directions of fundamental importance, namely, the resolution of the mass lines of the heavier elements and the measurement of the divergences from the whole number rule, a considerably more powerful instrument would be required. The Department of Scientific and Industrial Research, to whom I should like now to express my thanks, provided a liberal grant to defray the expenses of construction, and preliminary work on the instrument was commenced in 1921. The application of the method of accelerated anode rays led to an unexpected lengthening of the useful life of the original apparatus so that it was considered best to hold up construction of the new one in order that the final design might have the advantage of all accumulated experience. In the meanwhile one of its objects, the measurement of the divergences from the whole number rule, had been attacked by Costa† in Paris, using a mass-spectrograph of his own design capable of an accuracy of 1 in 3000. This admirable piece of work will be referred to later. The accurate determination of these divergences is of fundamental importance since it is one of the few avenues by which the problem of the structure of the nuclei of atoms can be

* F. W. Aston, 'Phil. Mag.,' vol. 49, p. 1192 (1925); also 'Isotopes,' 2nd edn., Arnold, London, 1924.

† J. L. Costa, 'Ann. Physique,' vol. 4, p. 425 (1925).

approached, and it was worth making every effort to push the accuracy of analysis to its extreme limit. It was finally decided that the increase of resolution could best be obtained by doubling the angles of electric and magnetic deflection, and sharpening the lines by the use of finer slits placed further apart, in addition special methods were considered for the necessary increase of accuracy in measurement. After numerous setbacks all these objects have been successfully carried out. The new instrument has five times the resolving power of the old one, far more than sufficient to separate the mass lines of the heaviest element known. Its accuracy is 1 in 10,000 which is just sufficient to give rough first order values of the divergences from whole numbers.

Apparatus.

The general appearance of the new instrument can be gathered from the photograph reproduced in fig. 1 and the details of its construction from the

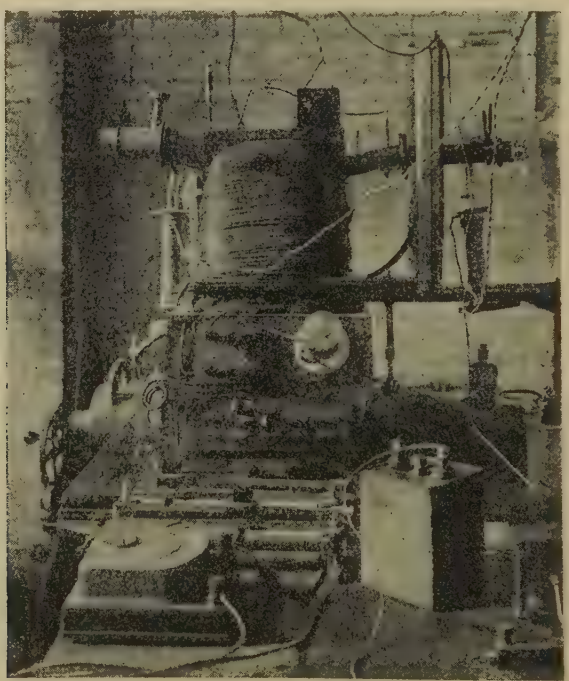


Fig. 1.

diagram, fig. 2, which is drawn approximately to scale. The large magnet was made by Messrs. W. G. Pye and Co., of Cambridge, who also carried out the difficult task of constructing the metal parts of the instrument. No

detailed drawings were available since the design of some portions could only be decided upon after others had been actually made, thus the best form of the

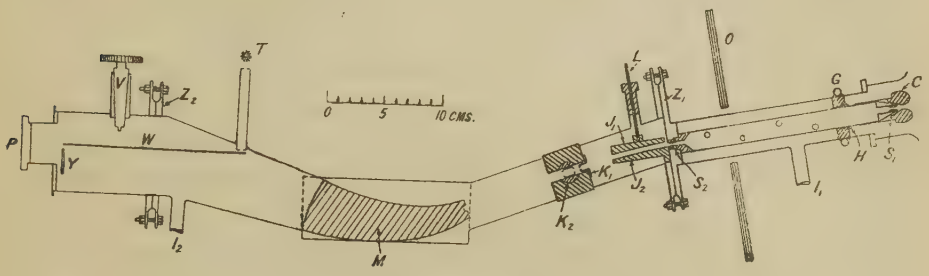


FIG. 2.

camera could only be arrived at by actual experiment with the rest of the apparatus in the laboratory. The excellence of the workmanship can be judged by the fact that the long ground joint by which the plate is moved after every exposure has been under full atmospheric pressure practically continuously for two years without even regreasing, and still shows no sign either of leak or tendency to stick.

The Discharge Tube.—Two types of discharge tube were employed during the preliminary work, the large bulb as used before, and a long cylinder with a tapered end of the form advocated by Wien and others. Conclusions as to the relative merits of these will doubtless be of interest to others working in this field. The bulb type was adopted in the early work on positive rays in the Cavendish Laboratory because it enabled the discharge to pass at the lowest pressures. This was of the first importance at a time when large "canal" tubes were used and no means, or only inadequate means, were available for reducing the pressure in the camera below that in the discharge. Nowadays the use of extremely fine apertures and improved means of maintaining high vacua have removed some of the objections to the higher discharge pressure in cylinders, and Wien claims that these are overwhelmingly superior to bulbs when the total quantity of rays passing through a small aperture at the centre of the cathode is measured. My own experience agrees with these measurements so far as they are concerned with the total quantity—as opposed to constitution—of the rays. I find that the bulb is in general inferior to the cylinder in this respect, but that if sufficient pains are taken to get the cathode into the best position this inferiority is not nearly so marked as Wien's figures suggest. The cylindrical tube is easier to make and to fit. It needs less care in adjustment and can be exhausted much quicker. It was therefore used for all testing

and preliminary work and proved of the greatest value in providing bright hydrogen lines for visual observation on the Willemite screen. On the other hand the rays produced by a cylindrical discharge tube have their energies concentrated into a very much smaller range than those from a bulb. In many experiments this is an advantage but in the type of measurement with which we are now concerned it is necessary to produce, at the same time, rays of energy ratio as high as 2 : 1. Furthermore the rays from a cylinder appear almost devoid of those multiply charged particles which are absolutely necessary to establish and check accurate ratios of mass. For these reasons the bulb type was used in practically all the experiments of high precision.

The concave cathode which is difficult to adjust, and also requires a special anticathode to dissipate the concentrated heat of its cathode rays, was replaced by one of the older convex type shown at C, fig. 2*, which though not so highly efficient is easy to adjust and has the merit of producing rays with a large range of energies. With this type of cathode 1 m.a. at 40 k.v. can be used for an indefinite period without any artificial cooling of the glass walls being required. The system of evacuation of the discharge bulb and of admission of gas to it was exactly the same as before ; the cathode and the wax joint behind it were kept cool by the circulation of water in the copper tube shown at G.

The Slit System.—The two slits S_1, S_2, which collimate the rays are of the same construction as before.† They are made as fine as possible consistent with reasonable exposures, the pair in use at present are about 0·02 mm. wide. Their distance apart is 20 cm., double that in the old instrument. They are mounted at the ends of a brass tube which also carries the cathode. This tube is supported at one end by its tip which fits accurately into a hole in the brass plate carrying the electric deflection system, and at the other by a ground joint H as indicated. By merely removing the discharge tube it is thus possible to examine or change both slits. A bayonet fitting ensures the carrier tube coming into exactly the same position every time it is replaced.

The Electric Field.—For the purpose in view it was desirable that this should be as efficient as possible since the higher the potentials necessary to deflect the rays through the large angle, one-sixth of a radian, designed, the more difficult they would be to produce, to apply and to measure. The plates were, therefore, curved and placed very close together. They were machined out of brass. The acting surfaces were ground accurately to 30 cm. radius,

* Sir J. J. Thomson, "Rays of Positive Electricity," Longmans, p. 30 (1921).

† F. W. Aston, 'Phil. Mag.,' vol. 38, p. 714 (1919).

1 cm. wide by 5 cm. long, and held 1·25 mm. apart by glass distance-pieces bearing on flats machined outside the curved faces. The lower earthed plate J_2 was bolted to the outer disc of the special joint Z_1 in such a way as to be capable of small adjustments both up and down and in rotation. The upper plate J_1 was clamped rigidly to the lower one by insulated screws and could be raised to the required potential through the insulated lead L. By the removal of this and the breaking of the special joint Z_1 and the connections for supply and exhaust of the discharge bulb the whole of the electric system could be removed for examination or adjustment. It can be easily calculated from these dimensions that the ratio of the energy of the ray in volts to the potential applied to the plates causing deflection through one-sixth of a radian is 120 to 1. This means that 400 volts is sufficient to deal with 48 kilovolt rays, about the hardest ever used. The potentials were derived from a set of 500 small accumulators built specially for the purpose. Very great care was given to the construction of these. They have been in use for five years and are charged twice a year only. After settling down they are remarkably constant. The change from day to day is barely perceptible, and there can be little doubt that the change in potential during the period of a single experiment is well under 1 in 10^5.

Special Joints.—It is clear that for the purpose in view rigidity of a high order is imperative, and for this reason the body of the apparatus was made of wide thick-walled brass tube permanently soldered together wherever possible. Demountable joints must be provided, however, and in order to secure vacuum tightness and rigidity these require careful design and construction. Costa* used ground joints. The type of joint, Z_1, Z_2, shown I find very convenient. It is easier to construct than a ground joint, and is of wider scope. It consists of two brass discs of liberal diameter turned with a small projection on their faces which keeps them concentric when locked together. The discs must be so thick that a deep groove, the wider the better, can be cut away without endangering the mechanical strength of the rims which are bolted together as shown. The mechanical strains are taken by the bolts while vacuum tightness is ensured by running a thin coat of soft Everett wax into the groove. When the joint is to be opened the seal is cracked cold, and when put together again it is usually only necessary to run a pea flame round the groove to ensure tightness. The most important point is for the bolts to be so far from the wax that they do not interfere with its accessibility to the flame or to the eye of the operator.

* *Loc. cit.*

The quantity of the electric spectrum of the rays allowed to pass into the camera was controlled by fixed and movable diaphragms K_1, K_2 exactly as before. The latter is mounted in a long and accurately ground taper plug of brass. The lever turning the plug is provided with a scale so that the aperture of the diaphragm can be set with great accuracy.

The Magnetic Field.—As before the instrument was built round its most massive member, the magnet. This was specially designed for the purpose. Its core was a ring of special magnet steel of rectangular cross section 15 cm. by 5 cm., with an external diameter of 46 cm. It was cut into three sections, a lower half and two upper quadrants. Between the upper ends of the latter a sector 5·2 cm. wide was removed to make place for the pole pieces of the mass-spectrograph. The three sections of the core were equipped with large brass end plates by which they could be bolted together. Since the principal object was to ensure an extremely constant field for considerable periods a minimum heating effect was desired both on account of this affecting the resistance of the windings and also the permeability of the core. This was obtained by winding the whole ring as lavishly as possible and so reducing to a minimum the watts per gram of copper necessary to produce a given field. By continuing the winding right up to the pole pieces a second valuable result is obtained, namely, the decrease of the stray field, the discharge is further protected from this by soft iron plates at 0. The three sections were, of course, wound separately. 225 lbs. of No. 14 D.C.C. wire were used, giving 6,257 turns in all. The magnet was mounted in a strong wooden cradle. It weighed about $4\frac{1}{2}$ cwt. and was somewhat bulky, but its performance has been very satisfactory.

The pole pieces form part of the walls of the mass-spectrograph itself into which they were soldered vacuum tight. Their rectangular outside section, indicated in the figure, registered with that of the core ring, and they were machined away smoothly so that the actual pole faces were of the greatly reduced sickle-shaped section shown at M. By suitable massive distance pieces these faces were held 3 mm. apart. At the point of entry of the bundle of rays small semi-circles were machined from the pole pieces in an attempt to achieve second order focussing,* but it is not yet certain whether any real advantage has been gained by this.

The length of the path of the rays in this highly-concentrated field is about 15 cm. The radius of curvature of the median ray is about 22·5 cm., so that the deflection of a singly-charged mercury atom with 30 kilovolts energy will

* F. W. Aston and R. H. Fowler, 'Phil. Mag.,' vol. 43, p. 523 (1922).

require a field of 15,700 gauss. On an actual test made by Messrs. Pye using the three windings in series with a total resistance of 17·9 ohms, the following figures were obtained :—

Current	0·5	1·0	2·0	5·0	amperes
Field	11,300	16,000	18,300	20,400	gauss

It will be seen that to deflect reasonably hard rays of the heaviest elements through two-thirds of a radian requires only about 10 per cent. of the technical safe current of the wire used. Below 10,000 gauss the field is practically proportional to the current, making it a simple matter to calculate the value of the latter necessary to bring a known mass to any required point of the spectrum. The current for the magnet is derived from a set of large storage cells and controlled by resistances. Standard resistances and a potentiometer are available if it becomes necessary to control the constancy of the current over very long periods. So far this has not been necessary for the effect of any small creep in the value of the magnetic field is eliminated by the particular methods of measurement employed.

The Camera.—The camera is made in two halves connected by a joint of the special kind described. The mechanism, not shown in the figure, for supporting the plate W, is mounted on the movable half so that it is only necessary to break one joint to enable any vertical adjustment to be made during the operations of focussing. Horizontal movement of the plate parallel to itself is achieved as before through a vacuum tight plug, V. The plate itself is a slip 2 cm. broad and 16 cm. long cut from a half-plate, and will easily accommodate six spectra. Paget half-tone plates have been used throughout the work. In a few trial cases these were Schumannised by the process already described,* but their enhanced resolving effect is not necessary with the new instrument and their increased sensitivity is more than counterbalanced by lack of uniformity.

A technical detail which has added greatly to the convenience of the operation of plate changing is the direct attachment of the plate glass window P to the tubular extension of the camera with Everett wax instead of a greased ground joint as before. The circular plate of thick glass is ground at the edge so that it projects slightly into the tube which it fits. The shoulder butted against the tube takes the strain of atmospheric pressure, and the small quantity of wax necessary runs naturally into the outer right angle corner giving a perfect lute. When it is required to change the plate, after air has been admitted to the

* F. W. Aston, ' Proc. Camb. Phil. Soc.,' vol. 22, p. 548 (1925).

camera, the wax joint is broken cold by a tap with a wooden pocket rule. The plate is changed and the window replaced at once as exactly as possible. The wax is resealed by running a small gas flame round the joint. The flame itself supplies the light necessary for the operation which only requires a few seconds and can be repeated scores of times with the same wax. The arrangement of the Willemite screen Y and the lamp T for recording a fiducial spot on the plate are exactly as before.

Vacuum Technique.—The method of obtaining the extremely low pressure in the slit system and camera has also been improved in many details. Coconut charcoal is still employed and is contained in pyrex glass tubes attached to the apparatus at I_1, I_2. Small electric furnaces, to be seen in the photograph, can be applied to these by which they can be kept at a definite temperature of about 400° C. for outgassing purposes. After a new plate has been put in the apparatus is exhausted at once by the rotating Gaede pump down to about 0·01 mm. with the furnaces on. The heating current is then cut off and the pumping continued for a short time as the charcoal cools. The stopcock leading to the pump is now turned off and the apparatus left for the night. Next morning the pressure is usually so low that no discharge at all will pass through the 8-inch bulb. Liquid air is now applied to the charcoal tubes and 2 to 4 hours allowed to elapse before gas is let into the bulb and the exposure started. No measurements of the pressure in the camera at this stage have been made, but it must be extremely small for perfectly sharp lines can be obtained although the rays travel about three-quarters of a metre before striking the plate. At the same time it looks as though the limit of the method is approached for a distinct falling off in sharpness is observed if stale liquid air is employed.

Progress of the Research.

It is not necessary to give the details of the lengthy process of fitting up the apparatus and adjusting the plate to the best focus. Except for one anomaly, the curvature of the slit images, which had never been satisfactorily accounted for in the first apparatus, and now appeared enormously exaggerated, progress was as expected. Even before the best focus had been found the resolution was sufficient to settle the constitution of mercury. But after the final adjustments had been made and the plate roughly calibrated an entirely unexpected difficulty arose.

The comparison of masses by the method of coincidence* depends on the theorem that if the magnetic field is kept constant and two mass spectra are

* See "Isotopes," 2nd Edition, p. 59.

taken with different electric fields, masses whose ratio is the inverse ratio of the electric fields will give lines coinciding exactly. This theorem had been tested by means of the lines of the atom and molecule of hydrogen using the bracketing method on the original mass-spectrograph, and found to give results correct to the accuracy then attained. When the atomic and molecular lines of hydrogen were tested in this way on the new apparatus a discordance was observed not merely measurable but absurdly large, several parts in 1000. The search for the cause of this was both long and disheartening for it must be remembered that the results of the smallest alterations in the apparatus take at least a day, generally several days to test. After months of disappointing work the conclusions reached were that the disturbing cause must be a charging up due to the impact of the rays, and its most probable location was the electric plates themselves. The fact that sharp lines could be obtained at all proved that it must have a steady maximum. Progress was at a complete impasse when Sir Ernest Rutherford kindly interested himself in the problem and suggested that an application of a rapidly reversed instead of a steady potential might show if the disturbing field took time to build up. By this means it was found that the cause was a polarisation of the surface of the plates which might take as long as 0·05 sec. to rise to its maximum value. Drastic scrubbing with emery paper reduced the effect temporarily and also the curvature of the lines which is due to the same cause, and the plates were later heavily gilded. With a clean gilded surface the effect was reduced to manageable proportions. Arrangements were now made to measure it with the highest accuracy possible. This was done by means of the hydrogen lines in the following manner. A special rotating contact breaker was made by which a voltage V, then a second voltage $V + h$ slightly different, and, finally, the two in series $2V + h$ could be applied rapidly over and over again. The length of the exposures at the first two potentials, giving lines of the atom, relative to the third, giving the line of the molecule, could be controlled and were adjusted to give lines of equal intensity. A ratio of precisely two to one should give three lines forming an exactly symmetrical bracket. This device eliminates the necessity of keeping the magnetic field absolutely constant since the electric fields are changed so rapidly that small changes will only lessen the sharpness of the lines and not alter their relative positions. It also makes it unnecessary to measure any potential at all so long as h is less than one per cent. of V. It is, in fact, the exact counterpart of finding the mid point of a line with a pair of dividers, and like this can, unfortunately, be applied only to ratios of 2 to 1. Measured by this means and allowing for the electron correction, the polarisation effect, even with clean

gold surfaces, was found to be still as high as 5×10^{-4}, much too high to be neglected. Experiments made while it was very much greater had, however, shown that it remained very constant so long as the conditions of discharge were not altered, so it was decided to eliminate it as far as possible by using the rays themselves to measure the ratios of the field strengths.

Performance of the Instrument.

The general appearance of the mass-spectra obtained by means of the new instrument in its present adjustment is shown by the reproductions on Plate 8. The spectrum photographed is nearly 16 cm. long and includes slightly more than one octave of mass. As Spectrum VI clearly shows, the distribution is surprisingly linear all along. The dispersion scale varies from about 1·5 mm. at the most deflected end, to rather more than 3 mm. at the least deflected end, for a change of mass of one per cent. Hence, in the middle portion of the plate the scale is about as open as that of a twenty inch slide rule. The resolving power is sufficient to separate lines differing by 1 in 600 and as expected from the theory is not very different at the two ends. Since the lines are irregularly curved and change their shape as one moves from one end of the spectrum to the other, it is impossible to assign positions to them relative to the fiducial spot with sufficient accuracy to approach the figure of 1 in 10^4 aimed at. This can only be done by measuring the distance between lines of approximately the same intensity, and therefore the same shape, when they are quite close together.

The Comparator.

Of several methods tried, the most convenient for making such a measurement is by an application of the principle of Poynting's Tilting Plate Micrometer.* If an object is observed through a thick plate of glass and the plate tilted, the object will appear to shift. Poynting showed that for moderately small angles, this shift is very exactly proportional to the tangent of the angle of tilt. For the present purpose, two plates are used with their edges touching. Their plane of separation is arranged to lie along the middle of the spectrum, which is observed directly below. The plates are so connected by a simple mechanism that they move through equal and opposite angles, giving double the range of a single plate, and affording complete optical compensation for focus, so that fairly high powers of magnification can be used when calibrating with a standard scale. To measure the distance between two neighbouring lines a and b, these are observed through a suitable low-power lens and the

* J. H. Poynting, 'Phil. Trans.,' A, vol. 204, p. 413 (1905).

plates tilted until the edge of the right half of *a* coincides with the edge of the left half of *b*, giving the appearance of a single line if the intensities are equal. The tangent reading is taken and the operation reversed until the edge of the left half of *a* coincides with the edge of the right half of *b*. The sum of these two readings from the central zero is a measure of the distance and can be calibrated once and for all by the use of a standard scale. The great advantage of this over other types of comparator, for this particular purpose, is that the setting is practically independent of small changes in the position of the eye of the observer. By a lateral movement of the eye one can shift the dividing plane right across the narrow spectrum, and so compare any part of one line with the corresponding part of the other. This enables much more reliable readings to be obtained with irregularly shaped but similar lines than would a fixed line of division. A diagram is shown in fig. 4, and

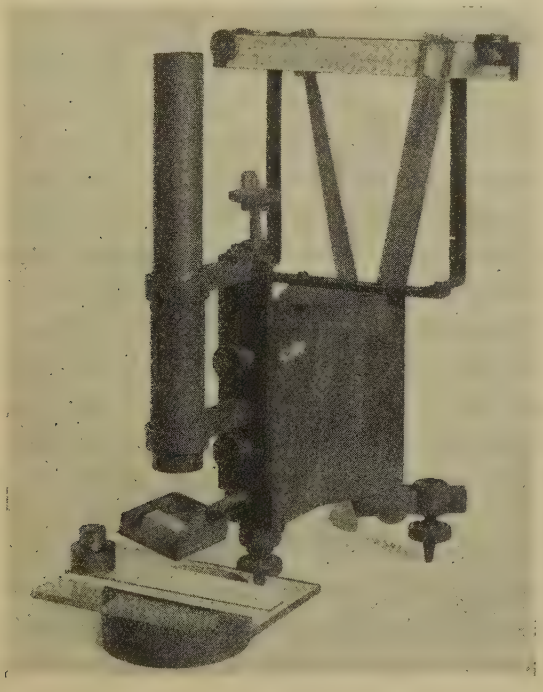


Fig. 3.

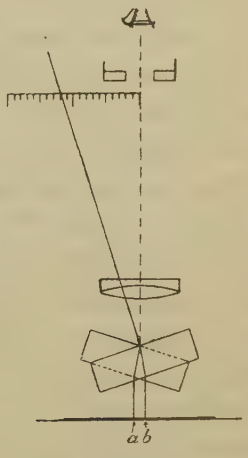


Fig. 4.

a photograph of the instrument itself in fig. 3. It is at present fitted with two echelon plates 1 cm. thick supplied by Messrs. Adam Hilger. The measuring arm is 30 cm. long, and the tangent scale divided into millimetres 78·5 of these correspond to 1 mm. shift. With sharp lines the accuracy of

setting is well within 0·01 mm., but the range is small, a little more than 2 mm. The fainter the lines on these spectra the sharper they are, so that the best measurements are to be obtained from lines just strong enough to be clearly seen. In order to increase contrast, the plates are laid face down on glazed process paper and observed by reflected light. In some cases the negatives were intensified before measurement. The position on the spectrum of the pair of lines to be measured is at our disposal, and is chosen to suit the particular system of measurement employed. All the most accurate determinations were made in that part of the spectrum where the dispersion constant is about 3 mm. for one per cent. change of mass.

Methods of measuring Ratios of Mass.

The accurate comparison of two masses giving lines on a mass-spectrum depends on the determination of two quantities, the distance between the lines, and the dispersion constant. The first must be small, and is measured directly by the comparator as described. The second is arrived at indirectly by approximate measurements of large intervals between lines whose mass differences are known with sufficient certainty. The hydrides of carbon and oxygen are most useful for this. From these data, the dispersion constant for several points on the small range actually used can be plotted. These lie on a straight line, so that interpolation is easy. Values of the dispersion constant appear reliable to about 0·3 per cent. They are only constant for one setting of the apparatus, and must be confirmed or recalculated if any change is made. It will be seen that if the interval between the lines corresponds to a difference of mass of less than 1 per cent., and no other errors are present, an accuracy of 1 in 30,000 can be looked for. These measurements can be applied in a number of different ways to suit particular cases, the principal ones so far used can be classified as follows :—

Method I.—Direct measurement on a single spectrum. This method is virtually free from all uncertainty and independent of the polarisation error, but can only be applied to bodies giving lines clearly resolved, but differing by less than 1 per cent. in mass. These cases are unfortunately very rare. The best example is the doublet given by oxygen, methane shown in Spectra I and IV. Here the lines only differ by 0·2 per cent. in mass, and can be obtained of equal intensity by manipulation of the quantities of oxygen and methane present in the discharge tube. Other cases will be noted later as they occur.

Method II, which may be called the method of series shift, can be employed whenever the masses to be compared form terms in a series whose unit of

difference is not too great. Two potentials are chosen which will bring consecutive terms into the desired contiguity, and these are applied to the electric plates alternately during the exposure while the magnetic field is kept constant. The relative duration of the application of the potentials is adjusted to give, in the double spectrum produced, equal intensity to those two lines between which the interval is to be measured. A good example of this is afforded by the lines of bromine and its hydrides. These give a series 79, 80, 81, 82, shown in Spectrum V. First potentials 300 and 324 volts were applied alternately, giving four times longer exposure to the latter to compensate for the weakness of the HBr lines. In this way, Spectrum II*a* was produced on which the small intervals 79, 80 and 81, 82 could be measured. The same operation was then repeated with the relative exposures reversed, giving Spectrum II*b* and the interval 80, 81. Now, it is obvious that whatever the value of the dispersion constant, so long as its variation along the plate is smooth, if the four lines form a true progression, the interval 80, 81 must be the mean of the intervals 79, 80 and 81, 82, each of which correspond to the addition of one hydrogen atom. When, as in this case, it is not, the divergence from the mean so measured gives at once the relation between the masses Br^{78} and Br^{81} required.

Method III.—This is the method most generally applicable for which the instrument was designed. It is the original bracketing method modified by the use of small intervals instead of brackets. If we wish to find the ratio of an unknown mass x to a known mass a we must photograph their respective lines in virtual coincidence, that is to say near enough for exact determination of the interval as described above. This is done by applying suitable potentials V_1, V_2 whose ratio differs by about half a per cent. from the expected ratio $x : a$. As we cannot trust the deflecting fields to be exactly measurable by the potentials applied, on account of the polarising effect, two other masses must now be found b, and c, whose ratio is identical or nearly so with $x : a$. The small interval x, a is photographed at the most advantageous position on the plate by applying the potentials V_1, V_2 in rapid succession alternately keeping the magnetic field H_1 constant. Then using the same pair of potentials and a different magnetic field H_2 the small interval b, c is photographed on another spectrum. The magnetic field in this case is so chosen as to bring both intervals into the same region of the plate to diminish as far as possible errors due to the values of the dispersion constant. The intervals of length are now measured and transformed into intervals of mass by multiplying by the dispersion constants calculated for the mid-point of each interval. From the value of the interval b, c, so determined, and the known ratio of the masses concerned the effective

ratio of the fields can be exactly calculated and then applied, by means of the interval x, a, to determine the mass x. In practice it is not usually necessary to evaluate the ratio of the fields, for the masses concerned are all so near whole numbers that the divergence of x from a whole number is simply the difference between the mass intervals, corrected by additions or subtractions of the known divergences of the other masses used. If the agreement between the ratios $x : a$ and $b : c$ is only numerically approximate, and the approximation of suitable order, the voltage ratio is adjusted to lie between them and the sum of the mass intervals used. The accuracy in this case will not be so high. An example of this method of comparison will be seen in Spectra III and IV. The first is a double spectrum taken with methane using a potential ratio between $15 \div 12$ and $24 \div 19$. The second is taken on the same plate with the same potentials after the magnetic field had been increased and BF_3 introduced. The sum of the intervals seen at the right hand of the spectra gives the relation between carbon and fluorine.

Units.—The choice of a standard of mass is at our disposal. From a theoretical point of view the neutral hydrogen atom, or the proton itself, would be a good unit, and would make all the divergences of the same, negative, sign. On the other hand, the fact that such masses as these lie at the extreme end of the scale makes them inconvenient as practical standards. For the present enquiry the neutral oxygen atom O^{16} has been adopted as standard. The identity of this scale with that of chemical atomic weights depends on whether oxygen is a simple element or not. The absence of a very small percentage of an isotope is difficult to prove, and in oxygen particularly so, for the neighbouring units 14, 15, 17, 18 are always liable to be present. The possibilities of an isotope O^{17} is actually suggested by Blackett's* experiments on the disintegration of nitrogen nuclei by the impact of alpha rays, but the evidence on the whole so far is in favour of oxygen being simple.

The masses measured by the mass-spectrograph are those of positively charged particles, and must, therefore, be corrected for the mass of the electron m_0 when this is significant. For this purpose m_0 is taken to be $0{\cdot}00054$ on the oxygen scale. To avoid ambiguity the word "mass" will always be used when the weight of an individual atom is concerned, "atomic weight" being given its usual significance. Where molecules are concerned their masses are assumed to be the exact sum of the masses of their component atoms.

* P. M. S. Blackett, 'Roy. Soc. Proc.,' A, vol. 107, p. 349 (1925).

The Packing Fraction.

Ever since the discovery of the whole number rule it has been assumed that in the structure of atoms only two entities are ultimately concerned, the proton and the electron. If the additive law of mass mentioned above was as true when an atomic nucleus is built of protons plus electrons as when a neutral atom is built of nucleus plus electrons, or a molecule of atoms plus atoms the divergences from the whole number rule would be too small to be significant, and, since a neutral hydrogen atom is one proton plus one electron, the masses of all atoms would be whole numbers on the scale H = 1. The measurements made with the first mass-spectrograph were sufficiently accurate to show that this was not true. The theoretical reason adduced for this failure of the additive law is that, inside the nucleus, the protons and electrons are packed so closely together that their electromagnetic fields interfere and a certain fraction of the combined mass is destroyed, whereas outside the nucleus the distances between the charges are too great for this to happen. The mass destroyed corresponds to energy released, analogous to the heat of formation of a chemical compound, the greater this is the more tightly are the component charges bound together and the more stable is the nucleus formed. It is for this reason that measurements of this loss of mass are of such fundamental importance, for by them we may learn something of the actual structure of the nucleus, the atomic number and the mass number being only concerned with the numbers of protons and electrons employed in its formation.

The most convenient and informative expression for the divergences of an atom from the whole number rule is the actual divergence divided by its mass number. This is the mean gain or loss of mass per proton when the nuclear packing is changed from that of oxygen to that of the atom in question. It will be called the " packing fraction " of the atom and expressed in parts per 10,000. Put in another way, if we suppose the whole numbers and the masses of the atoms to be plotted on a uniform logarithmic scale such that every decimetre equals a change of one per cent., then the packing fractions are the distances, expressed in millimetres, between the masses and the whole numbers.

Results.

The results obtained with the new instrument and now to be recorded may be classified under two entirely different heads. First there are those giving new information on the isotopic constitution of elements, and secondly there are those by which the packing fractions of the individual types of atoms are

measured. It is convenient to combine both of these under the element concerned, and, for ease of reference, to take the elements in their natural order of atomic number.

Hydrogen.—The hydrogen molecule was compared with the helium atom by Method III and measured against the known ratio $H : H_2$. The voltages applied were approximately in the ratio 2 : 1·004, so that the H_2 line was on the heavy side of each doublet. The difference between the packing fractions of hydrogen and helium is the sum of the two intervals corrected for the mass of the electron. The intervals of mass came out on three plates to be 73·7, 73·6, 73·9, mean 73·73. From this must be subtracted the correction for the electron which in this particular case amounts to $\frac{1}{4}\, m_0 = 1{\cdot}35 \times 10^{-4}$, whence we get 72·4 as the excess of the packing fraction of hydrogen over that of helium. The value of the latter is shown below to be 5·4, hence the packing fraction of hydrogen is 77·8, and therefore its mass 1·00778, a value in excellent agreement with the best results obtained by other means.

Helium.—The atom was compared with the doubly charged oxygen atom O^{++} using the known ratio $C^{++} : C^{+}$ as a measure. For this purpose voltages roughly 242 and 362 were applied to the plates, bringing C^{++} and He into close approximation on one spectrum and C^{+} and O^{++} together on the other. The packing fraction of helium will be measured by the difference between these intervals. The mean of four measurements gave 5·2. This must be corrected by the addition of $m_0/24$ so that the packing fraction of helium is 5·4 and its mass 4·00216, a value rather higher than 4·000 found by Baxter and Starkweather.*

Boron.—As before, boron trifluoride was found a convenient source of this element. The lighter isotope B^{10} was compared with O^{++} by the use of the known ratio $CH_3 : C$. B^{11} was compared with C by the known ratio $C : CH$, which is sufficiently near for the purpose. The results so obtained were checked by comparing the ratio $B^{10} : B^{11}$ with that of $B^{11} : C$. Using the mass of C given below, the results of these three comparisons were in good agreement, and gave for B^{10} the packing fraction 13·5, mass 10·0135; and for B^{11} the packing fraction 10·0, mass 11·0110.

Carbon.—The accurate evaluation of this atom is of peculiar importance, for it and its compounds give the most valuable standard lines used. Its mass can be measured in two ways. The more direct is to make use of the geometrical progression $O : C : OH_2$. The technical objection to this is that the water line is only well developed when a new discharge tube is fitted, and then its intensity

* G. P. Baxter and H. W. Starkweather, 'Proc. Nat. Acad. Sci.,' vol. 12, p. 20 (1926).

is very difficult to gauge. On the other hand the comparison is very favourable, for the square of the unknown is involved and any uncertainty in the mass of hydrogen only enters in the second order. The mean of four experiments so far made on this series gives a difference between the intervals $C : HO_2$ and $O : C$ of 2·7. The water molecule has a packing fraction 8·7 and the correction of the electron is quite negligible. Hence the packing fraction of carbon is half the difference, that is 3·0 and its mass 13·0036. The mass of carbon can also be measured by means of the O, CH_4 doublet. Several measurements of this have been made both by the comparator and by means of a photometer. From these the most probable value of the molecular weight of methane is 16·0350, a figure practically the same as that deduced by Baume and Perrot* from its density. The molecular weight worked out from the values for carbon and hydrogen given above is 16·0347 an agreement warranting confidence in the methods employed.

Nitrogen.—Work with this element is unsatisfactory owing to the impossibility of disentangling its lines from those of carbon compounds, and at the same time making use of the latter for ratios. By the use of nitrogen and ethane in varying proportions a certain success was obtained, and the measurements indicated the most probable value of its packing fraction to be 6·0 giving a mass 14·008. More work will be required to confirm this figure, it is however in excellent agreement with the chemical results.

Fluorine.—During the experiments with boron trifluoride the fluorine line was measured by means of the approximate geometrical progression $C : CH_3$ $F : C_2$. Spectra III and IV are examples of the results ; the measurements gave negligibly small packing fractions to fluorine as often negative as positive. The packing fraction is best taken as zero and its mass as 19·000, which is in agreement with the conclusions of Moles.†

Neon.—The more abundant isotope Ne^{20} was compared with O by means of the ratio $CH_3 : C$. The mean of three very consistent measurements of these intervals indicated that Ne^{20} has a practically negligible packing fraction 0·2 and a mass 20·0004. Ne^{22} was compared with Ne^{20} by the ratio $CH : C$. The numerical difference between these ratios is too large for accurate work, neither were the plates good. The mean results, a packing fraction 2·2 and mass 22·0048 are only to be regarded as provisional.

Phosphorus.—The element was introduced in the form of phosphine which gives the lines P, PH, PH_2, PH_3. If plenty of carbon monoxide is present its

* 'Compt. Rend.,' vol. 148, p. 39 (1909).

† E. Moles and T. Baruecas, 'Jour. Chim. Phys.,' vol. 17, p. 539 (1919).

line will be practically unaffected by the presence of small quantities of SI^{28}, inevitably present and of mass so far unknown. So that the series CO : P : PH_3 can be employed to give values for phosphorus. If the pair CO, P is photographed between the slightly wider pair P, PH_3 as indicated in Spectrum VII the two intervals can both be measured on the same spectrum, though it is safer to bring them into the same position on the plate by a second exposure. It is hardly necessary to state that in the particular spectrum reproduced the lines 28, 31 are much too intense to be of any value for actual measurement. From the known values of H, C and O and the sum of the two intervals the packing fraction of P can be calculated. The mean of six consistent values corresponds to a packing fraction $-5{\cdot}6$ and therefore a mass 30·9825. No mass-spectrum has given the slightest reason for supposing that phosphorus is complex so that it seems probable that the chemical atomic weight 31·02 is too high.

Sulphur.—The chemical atomic weight of sulphur 32·06 is known with great certainty. When it was first investigated with the old mass-spectrograph the results were indecisive for the instrument was quite incapable of separating the line due to molecular oxygen from that of a sulphur atom of mass 32·06 if that existed. Lines at 33 and 34 were noted, but by analogy to OH and OH_2 such lines were to be expected. Hence it seemed safest to rest content with the certain conclusion that atoms of mass number 32 were present in preponderating proportion. As soon as the new instrument was found to be capable of resolving doublets such as oxygen, methane, the matter was put to a direct test.* Spectra were obtained under conditions such that both O_2 and S must have been present but the line 32 showed no signs of doubling. This proved that the atom of S^{32} could not have a mass 32·06 so that atoms of higher mass number must be present. Further tests showed that the faint companions 33, 34 (S); 49, 50 (SO); ,65, 66 (SO_2), etc., were present on all spectra, even when the conditions were such as to make the presence of hydrides very unlikely. The intensity relations of the lines were also strongly indicative of true isotopic character. As a final confirmation a negative mass-spectrum of SO_2 was taken by exposing for an hour with both fields reversed. Lines 32, 33, and 34 were all visible and again showed the same intensity relations. We may safely conclude that sulphur, like magnesium and silicon, the two elements of even atomic number which precede it, is a triple element. The lightest mass number is by far the most abundant in all these three complex elements. S^{34} appears to be about three times as abundant as S^{33}; the two together amounting to

* F. W. Aston, 'Nature,' vol. 117, p. 893 (1926).

about three per cent. of the whole. The three lines of sulphur can be seen in Spectrum VI. No measurements of the packing fractions of the atoms of sulphur have yet been made.

Chlorine.—Cl^{35} was measured in three different ways. First by the use of a mixture of $COCl_2$ and CO_2 the approximate series $CO : Cl^{35} : CO_2$ was tested. Then by the addition of argon which had already been measured, the ratio $Cl^{35} : CO$ was compared with $A^{++} : O$. Finally Cl^{35} was compared directly with PH_3 by Method II. The results of all were consistent and gave the most probable packing fraction as $-4{\cdot}8$ giving a mass $34{\cdot}983$. Cl^{37} was compared with CO by the approximate ratio $C : O$, and also $CH_3 : A^{++}$, finally checked by comparison with Cl^{35} by Method II. These gave for Cl^{37} a value $-5{\cdot}2$ and mass $36{\cdot}980$.

Argon.—The principal isotope was first measured against CO_2 by means of the approximate ratio $P : PH_3$. Better results were obtained from its second order line using the more exact ratio $A^{40} : O$, $CH_3 : C$. The mean value for the packing fraction is $-7{\cdot}2$ giving mass $39{\cdot}971$. Experiments with the light isotope A^{36} were troublesome owing to its extreme faintness. It was measured by the approximate series $A^{36} : A^{40} : CO_2$. The measurements gave a packing factor $-6{\cdot}6$ and a mass $35{\cdot}976$, but this needs confirmation. During these experiments the proportion of A^{36} to A^{40} present was estimated by the relative exposures necessary to produce lines of equal intensity. It appears that this cannot exceed 1 per cent., and that therefore the atomic weight of the complex element cannot be less than $39{\cdot}93$, a result supporting the value $39{\cdot}94$ given by Moles* as against $39{\cdot}91$ the international figure.

Arsenic.—This element, like phosphorus, gives a series of hydrides so that it can be compared with Br^{79} with great accuracy by Method II using a series shift of one unit. The interval AsH_3, Br^{79} being measured against As, AsH on the one side and Br^{79}, $Br^{79}H$ on the other. The difference between the packing fractions of arsenic and Br^{79} measured in this way is very small, $0{\cdot}2$, so that if we take bromine as below, that of arsenic is $8{\cdot}8$ and its mass $74{\cdot}934$.

Bromine.—Methylene bromide was used as before, and it is of interest to note that the third order line of Br^{79} was obtained for the first time, it can be seen clearly at $26{\cdot}3$ on Spectrum VI, that of Br^{81} being masked by C_2H_3. The second order line of Br^{81} at $40{\cdot}5$ was compared with CO_2 by means of the ratio $C : CH$ giving as the mean probable value of its packing fraction $-8{\cdot}6$ and mass $80{\cdot}926$. Br^{79} was then compared with Br^{81} by a series shift of one unit as described above under Method II. The mean difference shown was $0{\cdot}4$ giving $-9{\cdot}0$ as the packing fraction of Br^{79} corresponding to mass $78{\cdot}929$.

* E. Moles, ' Ber. d. Chem. Ges.,' vol. 60, p. 139 (1927).

Krypton.—The position of this complex element on the mass scale can be best fixed by comparing the second order line of its heaviest isotope at 43 with CO_2 by means of the known ratio $Br^{79} : Br^{81}$. The mean of measurements on three plates gave for its packing factor $-8{\cdot}2$, mass $85{\cdot}929$. This result was checked by a comparison of the first order line of Kr^{86} with the second order line of the lightest isotope of mercury (Spectrum VIII) by means of the known ratio $CH : CH_3$ which, taking the value of mercury given below, gives identical agreement. The five even isotopes of krypton form an ideal subject for the series shift of two units. The six spectra shown (Spectrum XI) are all double spectra taken with the same voltage ratio, but with exposures suited to particular pairs of lines. The measurements show to an order of accuracy probably well within 1 in 10^4 that the masses of the isotopes of krypton form a perfectly uniform series. The determination of the common difference cannot be made with the same certainty, but it appears slightly greater than two units. An experiment with a series shift of one unit showed that the odd isotope Kr^{83} fell into the same series. The packing fractions and masses of the isotopes of krypton are given in the table.

Tin.—The application of the higher accuracy of the new instrument to tin was of peculiar interest for the following reason. The lines of tin had been obtained from its tetramethide with the old mass-spectrograph and on the plate the line of the monomethide of its most abundant isotope $Sn^{120}CH_3$ could be seen. The mass number of this being 135, it should have appeared symmetrically placed between the lines 134 and 136 due to the isotopes of xenon present by accident. Actually the asymmetry was so marked that it was estimated that the difference of the packing effect in the nuclei of the isotopes of tin and xenon must amount to at least $0{\cdot}2$ of a unit of atomic weight. This seemed theoretically unlikely, but the evidence then available appeared, and on re-examination still appears, quite conclusive. Spectrum X *a*, *b* shows the comparison made on the new instrument. The lines of xenon 132, 134, 136 and of tin monomethide 133, 135 are widely resolved and form a perfectly uniform series. It is therefore quite certain that the conclusions from the former observations were erroneous. The error appears to be due to the accumulation of several effects, an unequal spreading of the xenon lines, which were stronger than the line at 135, and an optical illusion possibly accentuated by the properties of the Schumannised plate. It is a striking example of the danger of drawing any conclusions whatever from the apparent positions of lines not clearly resolved.

Measurements with the comparator show perfect symmetry in the brackets

so that allowance for the excess due to CH_3 makes the packing fraction of tin 2·0 less than that of xenon. Taking the value for xenon given below Sn^{120} has a packing fraction −7·3 and a mass 119·912. As far as measurements have been made among the tin lines themselves it appears that they all have the same packing fraction. The clearer photographs obtained with the new instrument show that tin is even more complex than was previously supposed. On Spectrum X *c* taken with a very long exposure, no less than eleven isotopes are visible. Their mass numbers and order of intensity are given in Table I. Sn^{121}, previously suspected, is isobaric with the lighter isotope of antimony. Of the three new ones Sn^{112} and Sn^{114} are isobaric with cadmium and Sn^{115} with indium. Tin is quite exceptional. It is the most complex element known. It is the only element having more than two isotopes of odd mass number and the first to show isobares of odd mass number.

Iodine.—The single line of iodine was compared directly with that of Xe^{123} and found integral with it within the error of experiment. Until more accurate measurements are available the packing fraction of iodine can be taken as the same as that of xenon, namely −5·1, giving a mass 126·932 identical with that obtained by chemical means.

Xenon.—By a fortunate chance this element can be compared with mercury with the highest certainty, for its line 134 can be photographed in its second order between the two third order mercury lines 66·6 and 67·3, without its position being seriously disturbed by that of the weak mercury isotope 201. The mean of a large number of measurements of this bracket give a packing fraction 6·1 less than that of mercury. Hence, taking the value of mercury given below Xe^{134} has a packing fraction −5·3 and a mass 133·929. During these experiments the two doubtful isotopes of xenon 124, 126 were noted whenever sufficient exposure was given, and are confirmed beyond any reasonable doubt. A critical comparison of the masses of the isotopes of xenon by Method II has not yet been carried out, but direct measurement of the intervals between them shows no notable deviation, and for the present they may be taken as having the same packing fraction.

Tungsten.—During the preliminary work, an attempt was made to obtain the mass lines of this heavy element by the use of its hexafluoride. A sample of this volatile liquid was very kindly sent to me by Prof. F. Swarts of Ghent. The vapour was successfully introduced into the discharge, but appeared to decompose with great rapidity. No trace of any line could be detected in the region of 184. The failure of this experiment was, to a certain extent, expected from the behaviour of the lighter elements of the same chemical group. On

the other hand, had it been successful, it would have held out a hope of investigating the isotopic constitution of uranium, a problem of fundamental importance in radioactivity.

Mercury.—The value of the increased resolving power of the new instrument was first demonstrated in connection with this element. On spectra obtained with the old mass-spectrograph, the lines of mercury appeared as an unresolved blurr, the composition of which could only be estimated. As it turns out, the first estimate which included mass number 197 and omitted 201 was wrong. The new instrument resolved the mercury lines with the greatest ease as can be seen in Spectra V, VIII, IX and X. The groups show clearly that the mass numbers of its most important constituents are six :—198 (4), 199 (5), 200 (7), 201 (3), 202 (10), 204 (2). The numbers in the brackets indicate very roughly the relative intensity of the lines. This result was obtained in time to be of assistance in connection with controversy aroused by the claims of transmutation of mercury into gold.* The somewhat surprising fact that mercury occurs in very small proportion in the distillate from coal tar was brought to my attention by Mr. W. Kirby of the South Metropolitan Gas Company. A sample from this most unusual source was tested but found identical in isotopic constitution with the normal element.† Mercury is in some respects the most important of all elements in work on the determination of the masses of atoms by their mass spectra. Not only are its rays exceptionally easy to produce, but its unique series of multiply charged groups are to be taken gratefully as divisions naturally inscribed on the mass scale which connect the highest atomic masses with the lower ones without fear of accumulative error. During the present work the group could be distinguished even down to the sixth order. The accurate comparison between mercury and oxygen is of the highest importance so that it may be used as a sub-standard. This has not yet been done satisfactorily. A measurement of its packing fraction was obtained by comparing CO_2 with triply charged Hg^{198} by means of the known ratio of the second and third order lines of mercury themselves. This is by no means easy to do, for the potentials need be set very accurately to enable the two mercury groups to be clearly resolved, for, at the best, their interpolated lines can only differ by 1 in 400. Only one plate giving a good result has been obtained. On this the two intervals were indistinguishable. Hence, mercury would have the same packing fraction as CO_2, namely $0{\cdot}8$, and the mass of Hg^{200} would be $200{\cdot}016$. Comparison of the mercury lines with one another fails to show

* F. W. Aston, 'Nature,' vol. 116, p. 208 (1925).

† F. W. Aston, 'Nature,' vol. 119, p. 489 (1927).

any difference between their mass intervals and unity on the oxygen scale. They may all be taken to have the same packing fraction.

Costa's Results.

The work of Costa already mentioned consisted in comparing the masses of some of the lighter elements by the method of bracketing. His mass-spectrograph gave measurements for which an accuracy of about 3 in 10,000 is claimed. Taking helium as exactly 4 he deduced as the mass of the hydrogen atom 1·0079, but by a numerical slip this result was obtained by applying the correction for the mass of the electron to the hydrogen atom alone and not to the molecule of hydrogen and the atom of helium which were the bodies actually used.* The ratio of these two from the actual measurements was 2·0148 : 4 which is identical within his experimental error with that found in the present work. His comparison of helium with carbon giving the ratio exactly 1 : 3 is also, within his error, in agreement with the results now found for these elements. One may therefore use his experimental results for the isotopes of lithium with confidence. For Li^6 he gives a value of 6·009 if He = 4. If we substitute the value now found for helium we deduce for Li^6 a packing fraction of 20 and a mass 6·012. For Li^7 he gives 7·012 if N = 14·008, this corresponds to a packing fraction 17.

Collected Results.

The results obtained in the research can now be tabulated in their two respective categories :—

Table I.

Element.	Atomic number.	Atomic weight.	Minimum number of isotopes.	Mass numbers of isotopes in order of intensity.
S	16	32·06	3	32, 33, 34
Sn	50	118·70	11	120, 118, 116, 124, 119, 117, 122, 121, 112, 114, 115.
Xe	54	130·2	9	129, 132, 131, 134, 136, 128, 130, 126, 124
Hg	80	200·6	6	202, 200, 199, 198, 201, 204

In Table I is given a list of the elements on whose isotopic constitution further light has been thrown during this research. This may be regarded as supplementary to the complete table of isotopes last published.† With this it forms a complete list of non-radioactive isotopes discovered up to the present.

* Dr. Costa has requested me to make this explanation here.—F. W. A.

† F. W. Aston, ' Phil. Mag.,' vol. 49, p. 1199 (1925).

Table II.

Atom.	Packing Fraction $\times 10^4$.	Mass O = 16.	Atom.	Packing Fraction $\times 10^4$.	Mass O = 16.
H	$77{\cdot}8 \pm 1{\cdot}5$	1·00778	Cl^{35}	$-4{\cdot}8 \pm 1{\cdot}5$	34·983
He	$5{\cdot}4 \pm 1$	4·00216	A^{36}	$-6{\cdot}6 \pm 1{\cdot}5$	35·976
Li^6	$20{\cdot}0 \pm 3$	6·012	Cl^{37}	$-5{\cdot}0 \pm 1{\cdot}5$	36·980
Li^7	$17{\cdot}0 \pm 3$	7·012	A^{40}	$-7{\cdot}2 \pm 1$	39·971
B^{10}	$13{\cdot}5 \pm 1{\cdot}5$	10·0135	As	$-8{\cdot}8 \pm 1{\cdot}5$	74·934
B^{11}	$10{\cdot}0 \pm 1{\cdot}5$	11·0110	Kr^{78}	$-9{\cdot}4 \pm 2$	77·926
C	$3{\cdot}0 \pm 1$	12·0036	Br^{79}	$-9{\cdot}0 \pm 1{\cdot}5$	78·929
N	$5{\cdot}7 \pm 2$	14·008	Kr^{80}	$-9{\cdot}1 \pm 2$	79·926
O	0·0	16·0000	Kr^{81}	$-8{\cdot}6 \pm 1{\cdot}5$	80·926
F	$0{\cdot}0 \pm 1$	19·0000	Kr^{82}	$-8{\cdot}8 \pm 1{\cdot}5$	81·927
Ne^{20}	$0{\cdot}2 \pm 1$	20·0004	Kr^{83}	$-8{\cdot}7 \pm 1{\cdot}5$	82·927
Ne^{22}	(2·2 ?	22·0048)	Kr^{84}	$-8{\cdot}5 \pm 1{\cdot}5$	83·928
P	$-5{\cdot}6 \pm 1{\cdot}5$	30·9825	Kr^{86}	$-8{\cdot}2 \pm 1{\cdot}5$	85·929
			I	$-5{\cdot}3 \pm 2$	126·932
Tin (eleven isotopes)	....	....	Sn^{120}	$-7{\cdot}3 \pm 2$	119·912
Xenon (nine isotopes)	....	....	Xe^{134}	$-5{\cdot}3 \pm 2$	133·929
Mercury (six isotopes)	....	....	Hg^{200}	$+0{\cdot}8 \pm 2$	200·016

Table II gives a list of the precision measurements including those of Costa. The margin of error given may be regarded as an outside limit, hence it is large in all cases where comparisons are made indirectly. The masses of the atoms are simply calculations from the packing fractions so no particular significance is to be attached to the final digit. The agreement with chemical and other results is generally good in cases when such agreement is to be expected, but the best test of the accuracy of the measurements is their consistence among themselves. So far this is highly satisfactory, and makes it reasonably probable that most of the packing fractions are within one unit of their true value.

Conclusions.

As has already been explained, in addition to the first two fundamental constants of an atom, atomic number and mass number which settle the numbers of protons and electrons contained in its nucleus, we now have a third, the packing fraction, which gives entirely new information on the nucleus, for it is a measure of the forces binding those protons and electrons together. The discriminating value of this information is clear at once, *e.g.* had the packing fraction of the helium atom not been greater than that of the oxygen atom it would have ruled out the possibility that the nucleus of the latter was simply built of four unchanged helium nuclei or alpha particles, for there would have been no loss of energy, that is mass defect, in the latter to represent the binding

forces holding the four particles together. High packing fractions indicate looseness of packing, and therefore low stability, low packing fractions the reverse. We are at once led to inquire into what happens to the packing fraction as we ascend from atom to atom in the scale of mass.

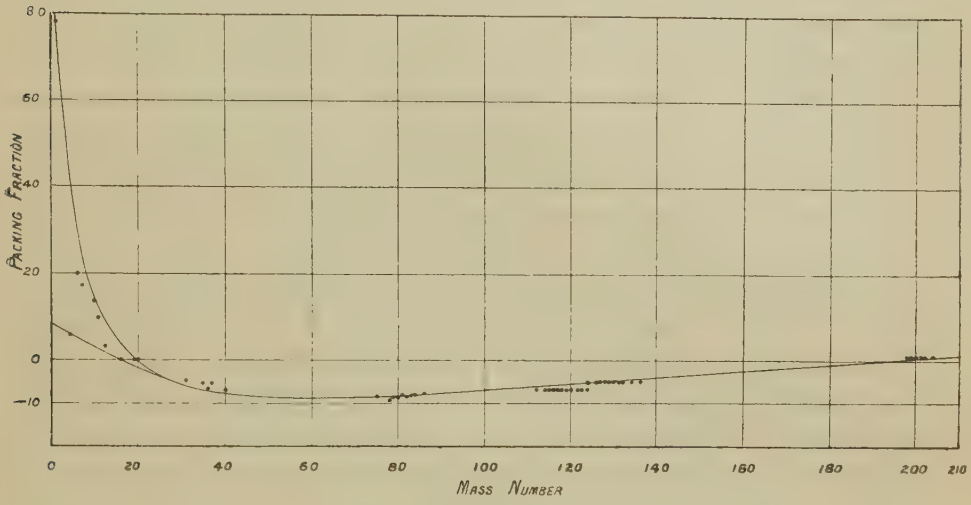


FIG. 5.

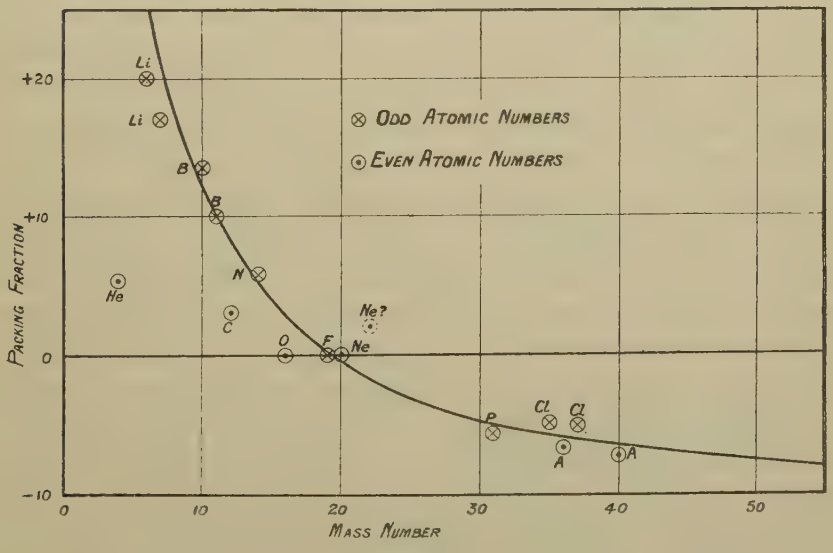


FIG. 6.

The result of plotting packing fractions against mass numbers for all atoms so far investigated is indicated in fig. 5. Fig. 6 gives the same plot for the

light elements on a larger scale. These indicate that the packing fraction as a function of mass number shows simple regularities of a remarkable kind. If we ignore for the present the large gaps which it is hoped to bridge as the work proceeds, it appears that all atoms, except those of light elements of even atomic number, approximate to a single curve. Starting at hydrogen with a large positive packing fraction the curve drops rapidly, crosses the zero line in the region of mass number 20 and sinks to a minimum value of about -9 in the region of mass number 80. It then rises again and recrosses the zero line in the region 200. There is no marked periodicity. As was to be expected, the most interesting region is that of the lightest atoms, shown in fig. 6. In isotopic constitution and relative abundance there is a fundamental class difference between elements of odd and elements of even atomic number.* This is reflected in the behaviour of their nuclei under the disintegrating impact of alpha particles, for Rutherford and Chadwick have shown that odd numbered elements have less stable nuclei and emit protons at a much higher average speed than those of even number. This difference is now shown to extend to their nuclear masses. Whereas the atoms of odd atomic number, irrespective of whether their mass number is odd or even, approximate to a smooth curve rising steeply to hydrogen, those of even atomic number lie well below and may be said to form a branch rising less steeply to helium. For comparison, the rectangular hyperbola $x\,(y+12{\cdot}5) = 250$ has been drawn in fig. 6. Such a curve is the locus of the packing fractions atoms of mass = mass number$+1/40$ unit, if $O = 16\,(1-12{\cdot}5\times10^{-4})$.

In the cases of atoms of odd atomic number the measurements show a definite approximation to this curve. This means that the masses of the nuclei of these particular atoms can be regarded as being made up of two distinct parts, one which changes by unity for each unit advance in mass number, and another, a small excess, which remains constant. To illustrate this point we may imagine the nuclei of these atoms to consist of a central core of maximum tightness of packing, corresponding to a packing fraction on the oxygen scale of $-12{\cdot}5$ which is surrounded by, let us say, three protons or neutrons attached with a tightness represented by a packing fraction $83{\cdot}3$. The free proton has a packing fraction on the new scale of $77{\cdot}8 + 12{\cdot}5 = 90{\cdot}3$ which leaves a balance of 7 for binding purposes. No stress is to be laid on these figures which are purely illustrative nor is it intended to discuss the possibilities in any detail here, but the facts do certainly suggest that the nuclei of light atoms have a loose, and therefore heavy, external structure of lightly bound protons or neutrons

* See 'Isotopes,' 2nd Edition, p. 131.

common to them but not possessed by the much more stable and tightly bound atoms of helium, carbon and oxygen Whether this is so or not the position of carbon and oxygen on the diagram indicates a tightness of packing entirely in favour of the views of Rutherford and Chadwick as against those of Kirsch and Pettersson who claim that protons can be detached from these nuclei by the impact of alpha rays. Neither do the observations of Rutherford and Chadwick that the chance of disruption of protons from lithium atoms, if possible at all, is small compared with those of boron, nitrogen, etc., show any serious discordance with Costa's results for lithium, here plotted, for the position of these suggests that their nuclear structure, though loose, is not so loose as that of the others. It is unfortunate in this connection that data for beryllium are not available. The fact that the packing fractions of the heavier atoms show a smooth distribution and do not decrease continually with increase of mass number is interesting. It is not what one would expect were the nucleus a structure similar to the outside of the atom, and possessing a periodic function. It is much more in keeping with the view put forward by Sir Ernest Rutherford.* That the nucleus consists of an inner part of uniform, tightly bound " crystalline " structure, outside which is a looser system of neutrons, protons and electrons which is more complex the heavier the element. The crystalline packing supplies a minimum possible packing fraction, while the increasingly complex outer structure may be taken to explain the rise of the packing fraction in very heavy elements.

The results reported here show that the mass-spectrograph has possibilities as an instrument of precision. The work is still in progress, and as it proceeds we may confidently expect an increase in the accuracy as well as in the numbers of measurements. It is hoped that these will form useful data for the theorist in the attack now imminent, on that least understood and most interesting problem of modern physics, the electromagnetic structure of atomic nuclei.

Summary.

A new mass-spectrograph is described with a resolving power of about 1 in 600, and an accuracy of 1 in 10,000, when special methods of comparison are employed.

By means of this instrument the isotopic constitution of mercury has been settled, new isotopes discovered in sulphur and tin and the two doubtful isotopes of xenon confirmed.

* Guthrie Lecture, 1927.

Measurements of precision have been made on 51 types of atom contained in 18 different elements ranging from hydrogen to mercury. The masses of these accurate to one or two parts per 10,000 are tabulated on the oxygen scale.

The relations of tin and xenon have been re-examined and found not to show the striking abnormality previously suggested.

Values for Li^6 and Li^7 are obtained by a re-calculation of Costa's results.

The packing fractions of the atoms, *i.e.*, their percentage divergence from whole numbers expressed in parts per 10,000 are plotted against their mass numbers when it is found that all but light atoms of even atomic number lie roughly on a single curve which descends from hydrogen 77·8 to a minimum of −9 in the region of bromine. It then ascends again and recrosses the zero line in the region of mercury.

The light atoms of even atomic number have packing fractions well below this curve and approximate to a branch rising much less steeply to helium 5·4.

The theoretical implications of these results are briefly discussed.

Description of the Mass-spectra reproduced in Plate 8.

I.—Single spectrum of the C_1 group, showing oxygen methane doublet.

II.—Double spectra illustrating comparison of Br lines by Method II.

III.—Double spectrum comparing CH_3 and C, voltages 280, 352. The oxygen methane doublet is seen on the extreme right.

IV.—Double spectrum comparing F and C_2, voltages 280, 352. This combined with III gives the mass of fluorine in terms of carbon and hydrogen by the general Method III. The line of B^{11} can be seen and, very faintly, that of B^{10}. The oxygen methane doublet is shown very clearly in the lower potential spectrum.

V.—Single spectrum taken with CH_3Br showing the pairs of lines due to Br, HBr, CH_2Br, CH_3Br and the second order mercury group.

VI.—Double spectrum comparing Br^{81} with CO_2, voltages 300, 324. The higher voltage spectrum, which had an exposure of half an hour, shows the triply-charged Br line at 26·3. It contains the lines of sulphur, chlorine, etc., and is an admirable illustration of the linear distribution of lines differing by one unit.

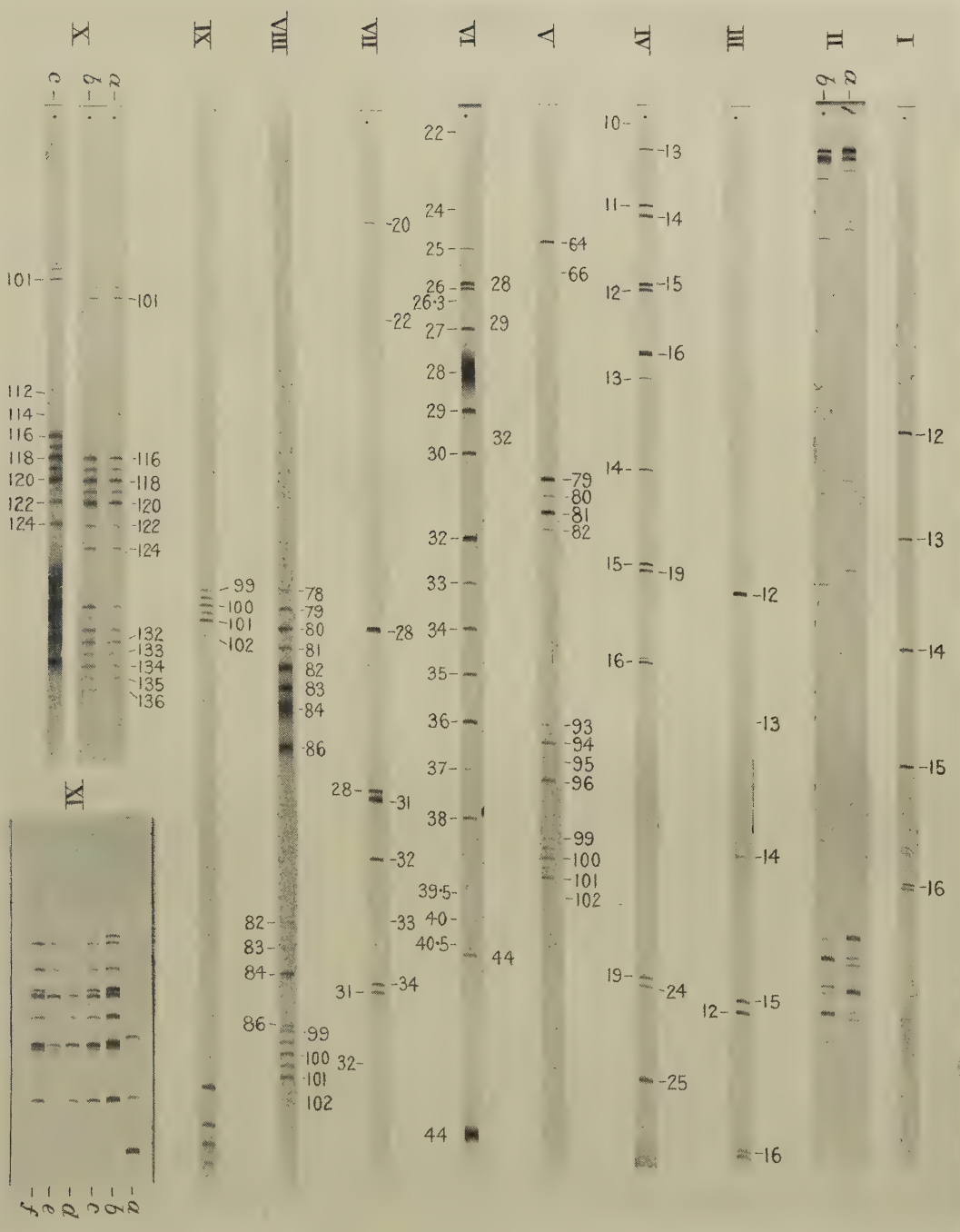
VII.—Double spectrum showing the lines of phosphorus and its hydrides photographed between the lines of carbon monoxide and phosphorus.

VIII.—Double spectrum comparing Kr^{86} with Hg^{198}. The lines of krypton are seen near the middle. The second order lines of mercury are shown under high dispersion.

IX.—Spectrum showing the second order mercury group very clearly.

X.—(*a*) and (*b*) Spectra showing the even spacing of the tin monomethide and xenon lines. (*c*) The same with long exposure showing eleven isotopes of tin.

XI.—Six double spectra taken with the same voltages 360, 368. Each has an exposure suited to a particular pair of lines, for comparison of CO_2, Kr and Br. (*a*) $Kr^{86} : CO_2$, (*b*) $Kr^{78} : KR^{80}$ and $Kr^{80} : Kr^{82}$, (*c*) $Kr^{80} : Kr^{82}$ and $Kr^{82} : {}^{84}$, (*d*) $Kr^{82} : Kr^{84}$, (*e*) $Kr^{84} : Kr^{86}$, (*f*) after addition of methylene bromide, $Br^{79} : Br^{81}$.



(*Facing p.* 514.)

On the Wave-Length of the Green Auroral Line in the Oxygen Spectrum.

By Prof. J. C. McLennan, F.R.S., and J. H. McLeod, M.A.

(Received June 15, 1927.)

[Plate 9.]

I. *Introduction.*

It is well known that the spectrum of the aurora is characterised by two outstanding features, the first of which is a set of bands with heads at or near $\lambda = 3914$ A, $\lambda = 4278$ A, and $\lambda = 4708$ A. The second is a strong narrow sharply-defined line close to $\lambda = 5577$ A. As to the bands, Lord Rayleigh,* Dr. Slipher,† Prof. Vegard‡ and others have shown them to be identical with the so-called " negative " bands obtained with molecular nitrogen in the singly-ionised state. Nitrogen in this state must, therefore, be one of the main constituents of that portion of the upper atmosphere in which auroral displays occur. As to the line $\lambda = 5577$ A, it is the strongest constituent of the spectrum of the aurora. Lord Rayleigh,* Dr. Slipher,† Dr. Babcock§ and others have shown that it can be obtained as well in the spectrum of the light of the night sky.

In 1923 Dr. Babcock§ made an accurate determination of its wave-length with a Fabry and Perot interferometer and found it to be $5577 \cdot 350 \pm 0 \cdot 005$ I.A.

In the same year Prof. Vegard|| put forward the view that the auroral green line had its origin in solid nitrogen suspended in a state of fine division in the upper atmosphere, and excited in some way to luminescence. In 1924 Vegard¶ as well as McLennan and Shrum** showed that the spectrum of the light from solid nitrogen rendered luminescent by electronic bombardment included three broad diffuse bands shading into each other with mean wave-lengths 5554 A, 5617 A, and 5658 A, but neither Prof. Vegard nor any other observer has obtained, as yet, from nitrogen in any form or state a spectral line possessing the characteristics of the auroral green line and having the wave-length $5577 \cdot 35$ A.

* Rayleigh, ' Roy. Soc. Proc.,' A, vol. 100, p. 114 (1922).
† Slipher, ' Astrophys. J.,' vol. 49, p. 266 (1919).
‡ Vegard, ' Phil. Mag.,' vol. 46, p. 193 (1923).
§ Babcock, ' Astrophys. J.,' vol. 57, p. 209 (1923).
|| Vegard, ' C. R.,' vol. 176, p. 1488 (1923).
¶ Vegard, ' C. R.,' vol. 180, p. 1084 (1925).
** McLennan and Shrum, ' Roy. Soc. Proc.,' A, vol. 106, p. 138 (1924).

On the other hand, McLennan and Shrum* in 1925 showed it was possible to obtain such a line in the spectrum of highly purified oxygen. When obtained from this gas alone the intensity of the line was weak. It could be enhanced, however, when the oxygen was mixed with one or other of the rare gases helium and neon.

In 1926 McLennan, McLeod and McQuarrie† published the results of an extensive investigation on the nature and occurrence of the line. They showed that argon in particular, when mixed with oxygen, enormously increased its intensity, and from observations made by them on the magnetic resolution of the line they were able to show that it was one that could be provided for in the scheme of spectral terms worked out for atomic oxygen by McLennan, McLay and Smith.‡ Apart from its wave-length, this interesting oxygen spectral line was shown to have all the physical characteristics of the famous green auroral line.

As to the wave-length of the green oxygen line, McLennan and Shrum* found it to be 5577·35 A with a possible error of 0·15 A. This value was the mean obtained from a number of settings on the line as recorded by them with an ordinary Hilger Constant Deviation Spectrograph. The possible error 0·15 A represents the greatest difference between their mean value and any one of the values they obtained from the different settings.

In view of the necessity of determining the wave-length of this oxygen green line with precision, we recently undertook to investigate it with a Fabry and Perot interferometer after the manner of Dr. Babcock.§ The investigation is now completed, and we find the wave-length to be $5577 \cdot 341 \pm 0 \cdot 004$ I.A.‖ As Babcock's value for the wave-length of the auroral green line was $5577 \cdot 350 \pm 0 \cdot 005$ I.A., it will be seen that there is a remarkable agreement between the results of the two determinations.

There would appear to be no doubt, therefore, that in the discovery by McLennan and Shrum of the oxygen green line the origin of the famous auroral spectrum line was found.

One result that follows immediately from the identification of the auroral green line is that oxygen as well as nitrogen must be present in the upper atmosphere in the region where the auroral light is emitted.

An account of our investigation follows.

* ' Roy. Soc. Proc.,' A, vol. 108, p. 501 (1925). † *Ibid.*, vol. 114, p. 1 (1927).

‡ *Ibid.*, vol. 112, p. 76 (1926). § Babcock, ' Astrophys. J.,' vol. 57, p. 209 (1923).

‖ Cario with a ruled grating recently found the wave-length to be 5577·348 ± 0·005 I.A. ' Z. f. Physik,' vol. 42, p. 15 (1927).

II. *Apparatus.*

The Fabry and Perot interferometer was a beautifully made instrument and was obtained from the Adam Hilger Company of London. The various parts are grouped and shown in fig. 1. The camera, a fixed focus one, was

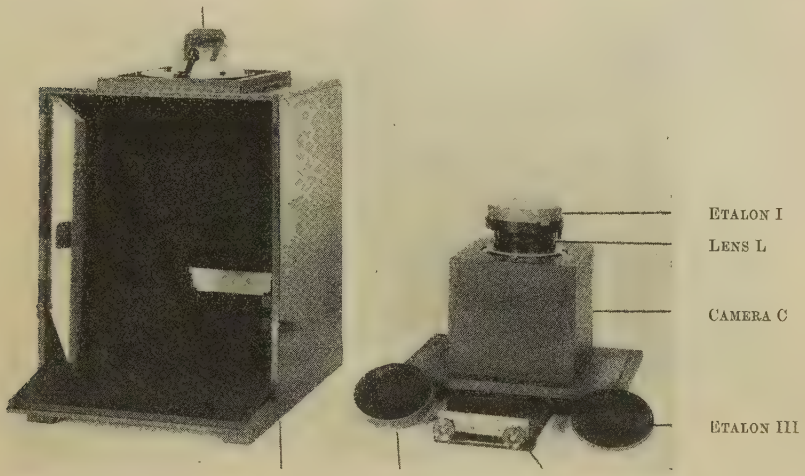


Fig. 1.

provided with an outer wooden enclosure A B on which a large right-angled reflecting prism, P, with its mount, could be placed. The face of this prism was 4·5 cm. × 4·5 cm., but the effective aperture was a circular opening in the base of the mount 3·5 cm. in diameter. C, the camera proper was provided with an F/2 Taylor-Hobson Cooke 4·5-inch focus anastigmat lens, L. On the mount of this lens there was a stage for holding securely when exposures were being made one or other of three Fabry and Perot glass etalons. The stage carried three levelling screws for adjusting the etalons as shown in fig. 3. The plate holder, G, partially inserted in the opening for it in the camera is shown in the photograph. The reflecting prism was used for directing a horizontal beam of light down into the camera. When a vertical beam was used the prism was removed.

The three etalons were plane parallel circular glass plates half silvered on both sides. Their diameters were the same, 7 cms., and their thicknesses, as will be seen later, were as follows :—

Etalon I = 1·57031 cm.
,, II = 1·20146 ,,
,, III = 0·46244 ,,

The following optical data were supplied by the makers :—

A.

Etalon.	Order of interference for $\lambda = 5577{\cdot}35$ A.	Spectrum range.*
I	85000	0·0656 A
II	65000	0·0858 A
III	25000	0·223 A

* Spectrum range in this table signifies the value of $\Delta\lambda$ that would result in the nth fringe of the $\lambda + \Delta\lambda$ system coinciding with the $(n + 1)$th fringe of the λ system.

B.—Refractive indices of the glass of the etalons. Melting No. 5358, type B.S.C.

μ_c (Hα, $\lambda = 6562{\cdot}793$) $= 1{\cdot}50627$
μ_d (NaD, mean, $\lambda = 5892{\cdot}95$) $= 1{\cdot}5086$
μ_f (H$_\beta$, $\lambda = 4861{\cdot}327$) $= 1{\cdot}51419$
μ_g (H$_\gamma$, $\lambda = 4340{\cdot}466$) $= 1{\cdot}51858$

The dispersion formula we used was

$$\mu = \mu_0 + \frac{C}{\lambda - \lambda_0} \text{ with } \mu_0 = 1{\cdot}4910853$$

$$C = 75{\cdot}358337 \text{ and } \lambda_0 = 1599{\cdot}17.$$

This formula gave $\mu = 1{\cdot}5100282$ for $\lambda = 5577{\cdot}35$ A and $\mu = 1{\cdot}5099835$ for $\lambda = 5586{\cdot}764$ A (Fe).

The manner in which the interferometer was used will be clear from the diagrams shown in figs. 2 and 3.

The light from the discharge tube in which the radiation corresponding to the oxygen green line originated was focussed by an astigmatic lens A on the slit B of a Hilger Constant Deviation Spectrograph. The eyepiece generally used in the spectrograph was removed and in its place a curved slit of adjustable width was placed at the focus, C. The curvature of this slit was made the same as that of the spectral lines in the region near λ 5577 A, so that while all the light in a selected spectral line could be passed through the slit, no extraneous light such as that from other spectral lines could get through.

Instead of allowing the beam of light diverging from the slit to go directly to the camera and so, by illuminating a wide area of the plate, form a complete ring system, we condensed it with a lens D so as to limit the illumination on the photograph to a narrow strip along a diameter of the ring system. This device served to reduce the time of exposure for the light of the oxygen green line from about 20 hours to 1 hour.

In our investigation we used as standards of comparison certain wave-lengths in the spectrum of the iron arc. The arrangement for passing the light from this

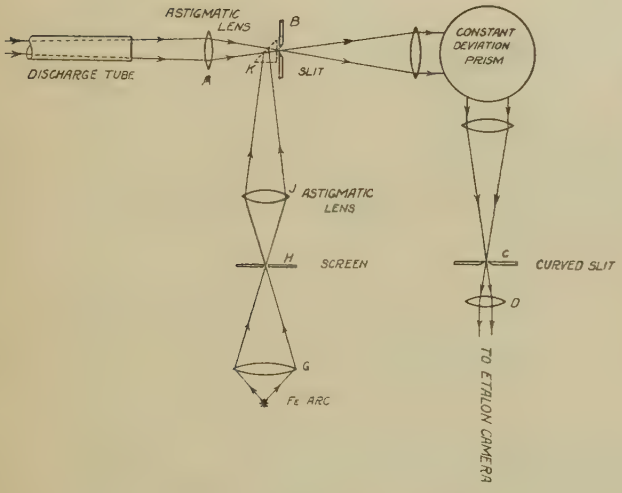


FIG. 2.

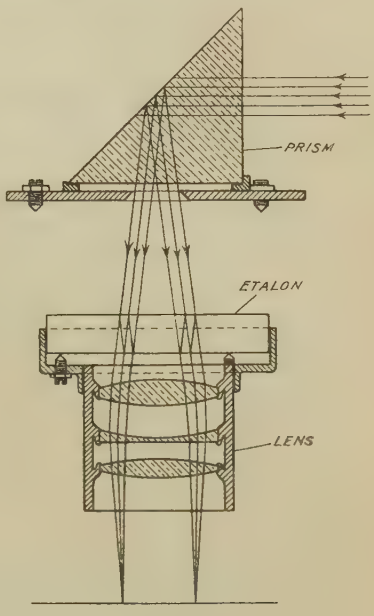


FIG. 3.

arc through the optical system, including the etalons, is also shown in fig. 2. The lens G formed a real image of the iron arc on the screen H, and an opening in this screen permitted the light from a selected portion of the arc to pass through the lens J and prism K.

When light from the iron arc was desired the prism K was placed in the position indicated, but it was removed whenever light from the oxygen discharge tube was under examination. Care was taken to ensure that the light from either source filled the lens of the collimator of the spectrograph in order to ensure symmetrical illumination.

A 6 mm. 6 ampere standard arc was tried as a source of the iron spectrum, but the spectral lines obtained with it were found to be too broad for use when the etalon with the highest order of interference was used. Accordingly a 12 mm. Pfund arc was substituted that carried a current of 3·5 amperes. Light was selected from a central portion of the arc 1·25 mm. in length. The iron lines obtained in this way were found to be fine enough to give well-defined interference rings with No. I, the 85000 order etalon.

III. *The Theory of the Wave-length Determination.*

For rays incident on an etalon at the angle i the equivalent path difference Δ_i in air is given by

$$\Delta_i = 2\mu e \cos i, \tag{1}$$

where μ is the index of refraction of the glass and e is the thickness of the etalon.

For normal incidence

$$\Delta = 2\mu e \tag{2}$$

For two wave-lengths λ_x and λ incident normally Δ_x and Δ the equivalent path differences are $\Delta_x = 2\mu_x e$ and $\Delta = 2\mu e$.

The order of interference for normal incidence for λ_x is given by $n_x + \alpha_x = 2\mu_x e/\lambda_x$ where n_x is integral. For λ we have similarly $n + \alpha = 2\mu e/\lambda$ where n is integral.

The difference in the order of interference for the two wave-lengths for normal incidence is given by

$$m + \delta = 2e\left(\frac{\mu_x}{\lambda_x} - \frac{\mu}{\lambda}\right) \tag{3}$$

where $m + \delta = (n_x + \alpha_x) - (n + \alpha)$ and m is integral.

It is not necessary to know the values of n_x and n in order to determine the value of δ, for $\delta = \alpha_x - \alpha$ when α_x is greater than α and $\delta = (1 + \alpha_x) - \alpha$ when α_x is less than α.

Equation (3) may be written in the form $\mu_x \nu_x - \mu\nu = (m + \delta)/2e$ where $\nu = 1/\lambda$.

From which

$$\nu_x - \frac{\mu}{\mu_x}\nu = \frac{m + \delta}{2\mu_x e}$$

$$\nu_x - \nu + \frac{\Delta\mu}{\mu_x}\nu = \frac{m + \delta}{2\mu_x e}$$

$$\Delta\nu - \frac{1}{\mu_x}\left(\frac{m + \delta}{2e} - \nu\Delta\mu\right). \tag{4}$$

By the use of this equation we may check up the value of $2e$, the double thickness of the etalon, and also, when the latter has been accurately found, evaluate the difference between the frequency of a radiation of unknown wave-length and that of a radiation whose wave-length is known with sufficient accuracy for its adoption as a standard.

Thickness of the Etalon.—To determine the thickness of the etalon with precision, it is first of all necessary to measure it directly with a gauge or some

similar instrument. A value can easily be found in this way correct to within one part in ten thousand.

For any two standard wave-lengths λ and λ_x we have

$$\mu = \mu_0 + \frac{C}{\lambda - \lambda_0} \quad \text{and} \quad \mu_x = \mu_0 + \frac{C}{\lambda_x - \lambda_0},$$

therefore

$$\Delta\mu = \frac{C(\lambda - \lambda_x)}{(\lambda_x - \lambda_0)(\lambda - \lambda_0)}.$$

Again

$$\Delta\nu = \left(\frac{1}{\lambda} - \frac{1}{\lambda_x}\right) \quad \text{since} \quad \nu = \frac{1}{\lambda} \quad \text{and} \quad \nu_x = \frac{1}{\lambda_x}.$$

All of which shows that $m + \delta$ in equation (4) can be determined if the values of λ and λ_x are known with the precision of standard wave-lengths.

While this procedure enables us to evaluate the quantity $m + \delta$, the important feature of it is that it gives us a method of precision for finding m the integral part of $m + \delta$. To find δ with precision an interference ring system should be photographed for the wave-length λ and also one for the wave-length λ_x with the etalon under investigation. If d_f and d_k be the measured diameters of any two rings f and k of the system formed by the wave-length λ, then α the fraction of an interference order at the centre greater than that corresponding to the first ring is given by

$$\alpha = \frac{(f-1)\,d_k^2 - (k-1)\,d_f^2}{d_f^2 - d_k^2}. \tag{5}$$

Similarly for the λ_x ring system we have

$$\alpha_x = \frac{(f_x - 1)\,d_{kx}^2 - (k_x - 1)\,d_{fx}^2}{d_{fx}^2 - d_{kx}^2}. \tag{6}$$

As $\delta = \alpha_x - \alpha$ or $(1 + \alpha_x) - \alpha$ as previously explained, we see equations (5) and (6) enable us with the measurements involved in them to evaluate δ with precision. If now we add this value of δ to the integral value of m found in the manner indicated above, we obtain an adjusted value of $m + \delta$ that must have a high degree of accuracy.

To determine $2e$ for the etalon with precision it only remains to insert the adjusted value of $m + \delta$ in equation (4), *i.e.* $\Delta\nu = \frac{1}{\mu_x}\left(\frac{m+\delta}{2e} - \nu\Delta\mu\right)$, and from it determine $2e$, the quantities $\Delta\nu$, μ_x, ν, and $\Delta\mu$ being known from the two standard wave-lengths used and the dispersion formula for the glass of the etalons.

In measuring the thicknesses of the three etalons used in our investigation

and in checking them up in the manner just indicated, we obtained the results given below :—

Values of $2e$, the double etalon thickness.

——	Direct measurement.	Optical measurement.	Difference.
			Per cent.
Etalon No. III	0·9248	0·924883	0·01
Etalon No. II	2·4028	2·402906	0·004
Etalon No. I.....	3·1408	3·140425	0·01

For this work the standard wave-lengths used in pairs were those of the iron arc λ 5405·778 A, λ 5429·702 A, λ 5434·527 A, λ 5455·613 A, λ 5586·764 A and λ 5615·653 A.

IV. *Wave-length Determination.*

For the actual determination of the wave-length of the oxygen green line the wave-length used as standard for comparision was λ = 5586·764 I.A. of the iron arc. This wave-length is one that is known to undergo a considerable amount of shift due to pole effect when the standard 6 mm. 6 ampere iron arc is used as the source of the radiation. As the radiation from the Pfund arc is free from all pole effect and gives sharper lines we used it in our determination. St. John and Babcock* have given the value of the iron arc wave-length used by us for a standard as 5586·765 I.A., while Monk† and also Meggers and Kiess‡ give it as 5586·764 I.A. The latter value was the one we used in our calculations.

The procedure adopted in measuring the wave-length of the oxygen green line was first to use Etalon III and with it to photograph the interference rings produced by the light of the iron arc standard wave-length ; then to expose another plate to the green line radiation, and finally to expose once more a third plate to the iron arc standard radiation. The exposures for the light from the iron arc varied from 2 to 5 minutes, and the exposure for the green line radiation from 40 to 60 minutes. The latter radiation was obtained from a mixture of oxygen and argon, with the latter in excess. The discharge tube was 100 cms. long and 1·5 cms. in diameter, and was operated with about 50 milliamperes from a 50,000-volt 3 k.v.a. transformer. Shifts in the fringes owing to temperature changes were reduced to a minimum by maintaining the room in which the photographs were taken at a constant temperature not only for the

* St. John and Babcock, ' Astrophys. J.,' vol. 53, p. 260 (1921).

† Monk, ' Astrophys. J.,' vol. 62, p. 375 (1925).

‡ Meggers and Kiess, ' Bur. of Stand. Sci. Papers,' No. 478 (1924).

duration of the exposures, but also for a considerable time before making them. From measurements on the diameters of pairs of rings as demanded by equation (5) values for α were found from the two plates taken with the light of standard wave-length, and the mean of these two values was adopted for the subsequent calculation of the phase constant δ. A value for α_x was similarly found from measurements on the ring system of the oxygen green line plate, and from these values of α and α_x the constant δ was determined. In finding the integer m for insertion in equation (4), *i.e.*

$$\Delta\nu = \frac{1}{\mu_x}\left(\frac{m+\delta}{2e} - \nu\Delta\mu\right)$$

it was necessary to use values for $\Delta\nu$, μ_x and $\Delta\mu$ derived from a value of λ_x for which the possible error was less than the spectral range of the etalon used.

As this range for Etalon III was 0·223 A, it was quite permissible to use the value $\lambda = 5577{\cdot}35$ A for the evaluation of m since the possible error in the value was certainly not greater than 0·15, and in all probability was very much less. The procedure just indicated enabled us to evaluate $m + \delta$. From the number so obtained the integral part was the value of m required. The following typical calculations will serve to show how $\Delta\nu$, and therefore λ_x, was evaluated :—

Etalon III.

Wave-length.	Readings on rings.				(Diameter)^2	k.	f.	α.
	No.	Left.	Right.	Diameter.				
5586·764 A	1	62·741	60·209	2·532	6·411	1	2	0·723
						1	3	0·722
	2	63·436	59·530	3·906	15·265	1	4	0·719
						2	4	0·704
	3	63·947	59·031	4·916	24·167	—	—	—
	4	64·369	58·608	5·761	33·189	Mean α =		0·717
5577·35 A	1	64·949	62·504	2·445	5·978	1	2	0·658
						1	3	0·661
	2	65·674	61·792	3·882	15·070	1	4	0·667
						2	4	0·700
	3	66·194	61·290	4·904	24·049	—	—	—
	4	66·596	60·864	5·727	32·799	Mean α =		0·672
5586·764 A	1	64·272	61·798	2·474	6·121	1	2	0·697
						1	3	0·691
	2	64·978	61·118	3·860	14·915	1	4	0·692
						2	4	0·683
	3	65·489	60·608	4·881	23·824	—	—	—
	4	65·906	60·193	5·713	32·638	Mean α =		0·691

Calculation of δ—

α_x for $\lambda\,5577{\cdot}35 = 0{\cdot}672$.

Mean α for $\lambda\,5586{\cdot}764 = 0{\cdot}704$.

$\delta = (1 + 0{\cdot}672) - 0{\cdot}704$.

$\delta = 0{\cdot}968$

Calculation of m—

$$\Delta\nu = \frac{1}{\mu_x}\left(\frac{m+\delta}{2e} - \nu\Delta\mu\right),$$

$$\nu_x = \frac{10^8}{5577{\cdot}35} = 17929{\cdot}662,$$

$$\nu_{Fe} = \frac{10^8}{5586{\cdot}764} = 17899{\cdot}449,$$

$$\Delta\nu = 30{\cdot}213,$$

$$\mu_x = 1{\cdot}5100282,$$

$$\mu_{Fe} = 1{\cdot}5099835,$$

$$\Delta\mu = 0{\cdot}0000447,$$

$$2e = 0{\cdot}92488,$$

therefore

$$m + \delta = 42{\cdot}935,$$

therefore

$$m = 42.$$

Calculation of λ_x—

$$\Delta\nu = \frac{1}{\mu_x}\left(\frac{m+\delta}{2e} - \nu\Delta\mu\right),$$

$$\mu_x = \mu\,5577{\cdot}35 = 1{\cdot}5100282,$$

$$\mu_{Fe} = 1{\cdot}5099835,$$

$$\Delta\mu = 0{\cdot}0000447,$$

$$2e = 0{\cdot}92488,$$

$$\nu_{Fe} = 17899{\cdot}449,$$

$$m + \delta = 42{\cdot}968 \qquad (m = 42,\ \delta = 0{\cdot}968),$$

therefore

$$\Delta\nu = 30{\cdot}236,$$

therefore

$$\nu_x = 30{\cdot}236 + 17899{\cdot}449,$$

$$= 17929{\cdot}685,$$

therefore

$\lambda_x = 5577{\cdot}343$ the wave-length of the oxygen green line.

Having obtained the value $\lambda = 5577 \cdot 343$ A for the wave-length of the oxygen green line by the use of Etalon III, our next step was to follow exactly the same procedure with Etalon II in place of Etalon III.

From measurements on the photographs of the ring systems formed by the oxygen green line and by the radiation of the iron arc of $\lambda = 5586 \cdot 764$ A the value of the phase constant δ was now found to be $0 \cdot 666$.

Taking $\lambda_x = 5577 \cdot 343$ A and $\mu_x = 1 \cdot 5100283$, it was found that $m + \delta = 111 \cdot 692$.

From the result $m = 111$ and for the value of $m + \delta$ after adjustment we thus had

$$m + \delta = 111 \cdot 666.$$

By using this value and $\mu_x = 1 \cdot 5100283$, $2e = 2 \cdot 4029$, $\nu_{Fe} = 17899 \cdot 449$ and $\Delta\mu = 0 \cdot 0000448$, it follows that

$$\Delta\nu = 30 \cdot 245,$$

therefore

$$\nu_x = 30 \cdot 245 + 17899 \cdot 449,$$

$$= 17929 \cdot 694,$$

therefore

$$\lambda_x = 5577 \cdot 340 \text{ A}.$$

Repeating the procedure with Etalon I, the constant δ was found to be $= 0 \cdot 988$ with λ_x taken equal to 5577·34 A, $m + \delta = 145 \cdot 973$. This gave an adjusted value for $m + \delta$ of $145 \cdot 988$. Using this value and taking $2e = 3 \cdot 140425$ $\Delta\nu$ was found to be given by $\Delta\nu = 30 \cdot 254$, from which $\nu_x = 17929 \cdot 703$ and therefore $\lambda_x = 5577 \cdot 337$ A.

Three determinations of the wave-length of the oxygen green line were made in the manner just described with each of the etalons, and the values found are given in Table I.

When the probable error of the mean wave-length was calculated from the deviations of individual determinations a value less than $\pm 0 \cdot 001$ A was arrived at. However, in order to allow for the possibility of systematic errors, we are placing the probable error of our result at $\pm 0 \cdot 004$ A, so that our statement of the wave-length of the green line in the oxygen spectrum is $5577 \cdot 341 \pm 0 \cdot 004$ I.A.

Plates 9A and 9B are reproductions of photographs taken with the 25000 order etalon when the light was concentrated into a narrow strip across the centre of the ring system. Plate 9A shows the section of the ring system produced by the light of the oxygen green line. Plate 9B is a similar reproduction

of a section of the ring system produced by the light of the standard iron line λ 5586·764 A.

Table I.—Wave-length Determinations.

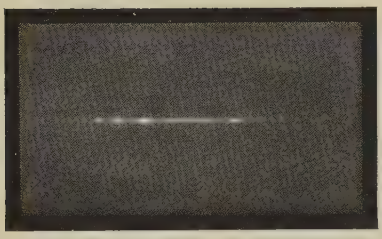
Oxygen green line by McLennan and McLeod.			Auroral green line by Babcock.	
Etalon.	Order of Interference.	Wave-length I.A.	Wave-length I.A.	Weight.
III	25000	5577·338	5577·348	2
III	25000	5577·343	5577·356	1
III	25000	5577·343	5577·353	1
II	65000	5577·340	5577·344	5
II	65000	5577·340	5577·346	3
II	65000	5577·337	5577·372	1
I	85000	5577·337	5577·355	1
I	85000	5577·345	5577·353	1
I	85000	5577·343	5577·350	1
			5577·352	2
			5577·352	2
Mean wave-length = 5577·341 I.A.			Weighted mean = 5577·350 I.A.	

In order to show the completed ring system, the lens D of fig. 2 was removed to allow the light from the slit to spread and illuminate a larger area of the photographic plate. The photograph reproduced in Plate 9c was taken in this way and shows the ring system for the oxygen green light. The total duration of the exposure was 18½ hours. Plate 9d shows the complete ring system produced by the light of the iron arc λ 5586·764 A.

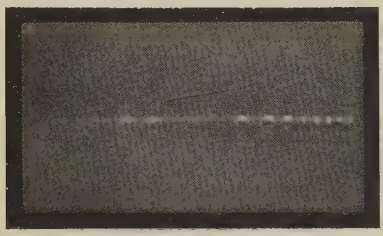
In fig. 4 there is reproduced a microphotometer graph of a plate taken of the fringe system for the green line radiation when the 85000 order etalon was

Fig. 4.

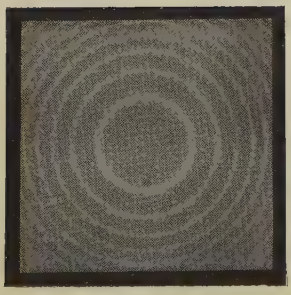
employed. It will be noticed that, even with this high order of interference, the rings were quite sharp. The spectrum range for the 85000 order etalon, it will be recalled, was 0·0656 A at λ = 5577 A. Measurements on the curve



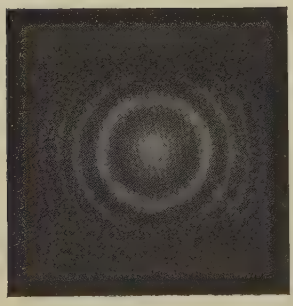
A.



B.



C.



D.

(*Facing p.* 526.)

of fig. 4 showed that the width of a ring was about 6/13 of the distance between successive rings, and therefore gave a value of 0·030 A as the approximate width of the oxygen line $\lambda = 5577{\cdot}341$ A under the conditions of our experiment.

V. *Conclusions.*

In conclusion, it is desired to draw attention to the remarkable way in which the present work confirms the view put forward by McLennan and Shrum,* namely, that the green auroral line has its origin in oxygen in the upper atmosphere.

The wave-length of the auroral line as given by Babcock† was 5577·350 I.A., and the value we got for the green line in oxygen was 5577·341 I.A. The difference between the two determinations of the wave-length is therefore only 0·009 A. It will be noticed from Table I, however, that in Babcock's list of individual determinations he considered the one which gave a value λ 5577·344 A to be the most reliable, and the one that gave the value λ 5577·346 A to be the second most reliable. The high weights placed on these particular values, namely, 5 and 3 respectively, may be taken to connote an even closer agreement between the two determinations than 0·009 would indicate. In any case, it would appear to us that the agreement in the wave-length determinations is amply sufficient to justify McLennan and Shrum's identification of the auroral green line with the green line of oxygen. Additional support for this view is found, moreover, in the fact that we have shown the width of the oxygen green line to be about 0·030 A. Babcock, it will be recalled, found the width of the auroral green line to be not greater than 0·035 A.

We wish to take this opportunity of acknowledging our indebtedness to the National Research Council of Canada for financial assistance in carrying out the investigation, and one of us, J. H. McLeod, wishes to express his indebtedness to the same Council for the grant of a Studentship which enabled him to take part in this research.

* McLennan and Shrum, 'Roy Soc. Proc.,' A., vol. 108, p. 501 (1925).

† Babcock, 'Astrophys. J.,' vol. 57, p. 209 (1923).

The Hydrogen Band Spectrum: New Band Systems in the Violet.

By O. W. Richardson, F.R.S., Yarrow Research Professor of the Royal Society.

(Received May 19, 1927.)

The system of bands described in this paper includes much of the strength of the secondary hydrogen spectrum when this is excited by direct electron impact on the H_2 molecule and there are no additional complications. I first observed it on a photograph of the spectrum of hydrogen at a pressure below 0·01 mm. of mercury, excited by a sharply limited electron stream,* which was kindly sent to me by Mr. P. M. S. Blackett. This photograph, which was taken by Mr. Blackett, had insufficient resolution and dispersion for the identification of the lines of the secondary spectrum. Up to the present I have not been able to get sufficient intensity at low pressures to obtain this spectrum with the purity shown on Mr. Blackett's photograph and the requisite resolution; but by working at pressures of about 0·3 mm. I have, with the help of Mr. D. B. Deodhar, got some plates on the large quartz Littrow spectrograph at King's College which show this spectrum sufficiently enhanced and the rest sufficiently reduced for the main features to be recognised with confidence.

The present paper deals with but a portion of what I have been able to arrange provisionally in this spectrum, but it is a unit which seems to be complete in itself, except that I am only able to give the Q branches of the bands, whereas there are indications, still insufficiently explored, that P and R branches do exist. There are two band systems, the first very strongly developed. The nucleus ($0 \rightarrow 0$ Q (1) line) of the first band is at $\lambda\,4633{\cdot}95\,(9) = \nu\,21573{\cdot}81$. The Q branch of the $1 \rightarrow 0$ band is the series 20Q (m) of Richardson and Tanaka.† All the bands are degraded towards the violet, *i.e.*, in the opposite sense to those in the systems described in Structure, Parts IV and V.‡ Each line of the bands is accompanied by a fainter component on the long-wave side of it. The intensity of the fainter line is usually about one-fourth of that of the stronger in terms of the numbers of the usual conventional scale of intensities, and the separation of the two lines is in the neighbourhood of 4 wave-numbers. The final states of the present bands appear to be the same as the initial states of the Lyman bands in the far ultra-violet.

* *Cf.* Blackett and Franck, 'Z. f. Physik,' vol. 35, p. 389 (1926).

† Richardson and Tanaka, 'Roy. Soc. Proc.,' A, vol. 107, p. 614 (1925).

‡ Richardson, 'Roy. Soc. Proc.,' A, vol. 111, p. 714; vol. 113, p. 368 (1926).

It will be convenient to refer to this system as the A bands, and to denote the $n' \rightarrow n''$ Q (m) line by the abbreviation $n'An''$ Q (m). The Q (1) lines are shown in such a way as to exhibit the vibrational changes in Diagram I. The lines in the same row have the same initial vibrational state n' and those in the same column the same final vibrational state n''. In each square the first number is the wave-number of the main Q (1) line of the band. Of the two numbers which follow this in brackets, the first is the intensity of the line as recorded by Merton and Barratt,* Tanaka,† or Deodhar.‡ The second is Deodhar's estimate of the intensity on a plate taken with the luminosity excited with a sharply limited stream of electrons. It will be convenient to refer to this discharge as the " field free " discharge, or f.f. discharge for short. On this plate the faint component is not in general resolved from the main line. The number below this gives the second difference of the wave-numbers of the Q (m) lines of the band. The next number is the wave-number of the fainter component of the

DIAGRAM I.

Q (1) and Q′ (1) Lines of A bands.

Initial \ Final	0	1	2	3
0	21573·81 (9) (10) 13·54 21570·50 (3) 12·95	20260·71 (4) (7) 19·74 20256·78 (0) 19·18	18984·34 (6) (1) 22·78 18980·52 (*r*) 22·79	
1	23933·69 (4) (10) 12·48 23929·91 (1) 12·82	22620·75 (1) (3) 15·68 22616·61 (1) 17·34	21344·26 (5) (7) 20·85 21340·39 (1) 21·23	20097·75 (rd^2) (?) 26·48 20093·59 (rd^2) 26·27
2	26261·75 (*p*) (*p*) 10·82 26261·75 (*p*)	24955·57 (1) (5) 15·85 24947·30 (*qd*) 12·82	23677·16 (3) (7) 20·15 23677·16 (3) 22·07	22431·48 (2) (4) 25·88 22427·41 (*p*) 26·51
3		27247·81 (*r*) (0) 14·94	25971·06 (0) (3) 16·53 25966·81 (*q*) 20·02	24723·65 (1) (5) 21·62 24718·03 (*p*)

* ' Phil. Trans.,' A, vol. 222, p. 369 (1922).

† ' Roy. Soc. Proc.,' A, vol. 108, p. 592 (1925).

‡ ' Roy. Soc. Proc.,' A, vol. 113, p. 420 (1926).

Q (1) line, followed, in brackets, by the intensity assigned to the line in the tables referred to above. It is convenient to indicate the fainter components by a dash, thus : ${}_0A_0'$ or ${}_0A_0$ Q′ (m) to indicate the fainter components of ${}_0A_0$ or ${}_0A_0$ Q (m) respectively. The small number below this is the second difference of the vibration numbers of the Q′ (m) lines.

The lines satisfy the combination rules that for any four adjacent squares the sum of the frequencies is the same for each diagonal. This applies both to the Q (1) and to the Q′ (1) lines and also to any similar diagram constructed with Q (m) or Q′ (m) lines with values of "m" other than 1. The lines are enhanced in the field-free discharge. The chief exception, 18984·34 (6) → (1), is much too strong for its place in the band and is thought to be mainly something else. The other, 26261·75 (p) → (p), is a faint line in a little explored part of the spectrum whose wave-number is much too low for the main line. It is probably either the wrong line or is mixed with something else. It is interesting to note also that ${}_1A_2$ Q (1) = 21344·26 is enhanced relatively less than the others (5) → (7). This line also does duty as ${}_3\beta_3$ Q (2), where it was pointed out* that its intensity was too high and its wave-number too low. Probably most of the strength of this line belongs to the present system. The most intense band of the system is the ${}_0A_0$ band, and the strength falls away in every direction from the left-hand top corner of Diagram I. Relative to the others the intensity of ${}_1A_1$ Q (m) seems low ; otherwise the intensity distribution among the different bands seems regular. The only line of any strength which does not possess a fainter component close to 3·7 wave-number on the long-wave side of it is ${}_2A_2$ Q (1) = 23677·16 (3) (7), and this is recorded by Merton and Barratt as an unresolved doublet. The second differences increase in proceeding horizontally from left to right in Diagram I and diminish in proceeding vertically from above downwards in the manner which is usual in vibration band systems when the bands are degraded towards the violet. The second differences in these Q branches are not constant, and there are some irregularities and missing lines in the weaker bands, so that any differences which may appear in the values of the second differences for the Q (m) and Q′ (m) branches for a given band are uncertain. It cannot be said that the differences of these second differences for the same band in Diagram I are greater than the uncertainties in the second differences themselves.

The lines are shown in Table I. In each square of this table the first number is the wave-number of the main line, followed, in brackets, by (1) the intensity

* Part V, *loc. cit.*, p. 383 ; Part IV, p. 733.

assigned to it in the tables of Merton and Barratt, Tanaka or Deodhar, and (2) the intensity assigned to them by Deodhar on the f.f. discharge plate already referred to. The letters T or D underneath indicate that the lines are to be found in Tanaka's or Deodhar's tables respectively. The other letters denote the character of the lines as given in Merton and Barratt's tables with the abbreviations H.P. = high pressure, L.P. = low pressure, C.D. = condensed discharge, He = helium effect, Z = Zeeman effect, and S = Stark effect. These are followed by claims made by other systems. Underneath, the same information is given with respect to the fainter constituent. The lines enclosed in rectangles are abnormally strong. Many of them are probably cases in which strength from other systems covers up a fainter member of the present bands, but others may prove to be genuine abnormalities. Most of them come up in the f.f. discharge.

In contrast to the $2\sigma - m\pi$ system many of the stronger lines of the present system show the Zeeman effect. It appears to be the vibrationless initial states which correspond to the magnetic field, but it is not universal. The strong lines of the ${}_1A_2$ band are specifically marked Z = 0 in Merton and Barratt's tables. In the ${}_1A_1$ band, which for some reason is weak compared with the bands which surround it in Diagram I, nearly all the lines are enhanced by helium. There are a fair number of interferences with lines of the $2\sigma - m\pi$ systems of Part V and with Allen and Sandeman's II and with various odd arrangements which I have proposed. In, I think, every case of interference with $2\sigma - m\pi$ the lines are too strong in one system or the other, generally in $2\sigma - m\pi$. Some of these cases have been noted in "Structure," Part V. In most of the other cases there is evidence that the interferences are genuine. Such interferences are likely to have a more serious effect on the wave-numbers of the weaker lines of higher rotation quantum number, and it is very satisfactory, as will appear below, that the term data got from the stronger Q (1) lines show an excellent mutual agreement.

The intensity of the lines of these bands exhibits a very pronounced alternation. This is well shown by an examination of the ${}_1A_0$ band, which is the most fully developed and extends to 7 lines. The intensities of these are :—

$m \rightarrow$	1	2	3	4	5	6	7
Int. (M. & B.)	4	2	3	1	2	0	0
Int. (f.f.d.)	10	4	8	2	6	2	3

The second differences of the Q branches are neither constant nor of the type which arises from a term in the energy which is proportional to the fourth power

of the rotational quantum number. Apart from some irregularities, they seem to be of the type shown by ${}_1A_0\,Q\,(m)$, in which the second difference first increases and then falls away as "m" is increased. The behaviour of this band (20 Q (m)) has already been described by Richardson and Tanaka.*

Table I.

	${}_0A_0$.	${}_0A_1$.	${}_0A_2$.
Q (1)	21573·81 (9)* (10)	20260·71 (4) (7)	18984·34 (6) (1)
	He+, Z	Z	He 0·154Q (7)
Q′ (1)	21570·50 (3)* (10)	20256·78 (0) (0)	18980·52 (r) (?)
	Z	L.P.++203P (4)	
Q (2)	21585·54 (3)* (10)	20276·93 (r) (a)	19007·25 (0) (a)
	Z	D	T
Q′ (2)	21583·45 (10)* (10)	20269·22 (3) (1)	19001·87 (4) (p)
	He+, Z		L.P.++He+
Q (3)	21613·76 (4) (4)	20314·80 (0) (r)	19051·36 (0) (a)
	HP++, He 0, Z . ${}_2\beta_2$ Q(3)	60R (1)	He+164R (5)
Q′ (3)	21606·53 (0) (?)	20309·56 (0) (?)	19046·24 (r) (a)
			D
Q (4)	21655·43 [–] (?a)	20369·09 (1) (?a)	19121·56 (p) (0)
	He++	L.P.++84Q (4)	T
Q′ (4)	21650·73 (3)* (8)	20365·89 (2) (1)	† 19116·92 (2) (0)
	L.P.++C.D.=${}_2\beta_2$Q(1)	L.P.++C.D.+He+104Q(2)	L.P.++He++
Q (5)	21707·79 (0) (1)	20439·11 (2) (1)	19209·38 (0) (p)
	H.P.+He+111P (5)	$3a_1$ Q (1)	$II_1 a$P (3)
Q′ (5)	21702·52 (r) (?)	20435·86 (q) (?)	19206·65 (0) (?)
	D	T	

* These lines are mixed with others on the f.f. discharge plate. Where these are in the same square the two lines are unresolved.
† Unresolved doublets (M. and B.).

* 'Roy. Soc. Proc.,' A, vol. 107, p. 614 (1925).

Table I—(continued).

	$^1A_0$.	$^1A_1$.	$^1A_2$.	$^1A_3$.
Q (1)	23933·69 (4) (10)	22620·75 (1)* (3)	21344·26 (5) (7)	20097·75 (rd^2) (?)
		C.D.+He+S	L.P.+($Z=0$) $_3\beta_3$ Q (2)	D
Q′ (1)	23929·91 (1) (?)	22616·61 (1)* (3)	21340·39 (1) (?)	20093·59 (rd^2) (?)
		L.P.++	H.P.+	D
Q (2)	23944·64 (2) (4)	22637·19 (1) (1)	21365·65 (3) (1)	20122·75 (r) (?)
		He+204R (5)	H.P.++ ($Z=0$)	D
Q′ (2)	23939·42 (r) (?)	22631·80 (3) (3)	21359·62 (2) (1)	20118·22 (3) (0)
	T $_2\beta_1$ Q (4)	He++, Z, S 112Q (3) 204R (1)	L.P.++	
Q (3)	23966·68 (3) (8)	22666·49 (2) (0)	21404·20 (0)* (6)	20173·01 (3) * (2)
			L.P.++	(He=0)
Q′ (3)	23958·64 (rd) (?)	22657·81 (4) (8)	21401·22 (3)* (6)	20168·77 (3)* (2)
	T II_c c Q (5)	He++, Z, S	H.P.++ ($Z=0$)	(He=0)
Q (4)	24001·88 (1) (2)	22716·74 (1) (3)	21468·35 (r) (a)	20253·00 (1) (?a)
		L.P.++He+	T	203P (2)
Q′ (4)	23996·47 (rd) (?)	22712·30 (0) (?)	21463·23 (1) (a)	20251·36 (1) (?a)
	D	He++	H.P.+	
Q (5)	24050·77 (2) (6)	22777·44 (1) (2)	21552·28 (0) (?)	20361·2–1–(?)
	He 0 . II_c b Q (5)	He++	60R (5)	T
Q′ (5)	24041·9—[—(a)	22777·44 (1) (2)	21546·15 (p) (?)	20356·44 (q) (?)
	T	He++	T	T
Q (6)	24111·55 (0)* (2)	22866·60 (p) (0)		
	S	T		
Q′ (6)	24106·78 (q) (?) T			
Q (7)	24182·92 (0) (3)	22963·96 (0)* (0)		
	S $_2r_2$ Q (2)			
Q′ (7)	24176·72 (r) (?)	22958·26 (0)* (0)		
	D	L.P.+		

* These lines are mixed with others on the f.f. discharge plate. Where these are in the same square the two lines are unresolved.

Table I—(continued).

	$^2A_0$	$^2A_1$	$^2A_2$.	$^2A_3$.
Q (1)	26261·75 (*p*) (*p*)	24955·57 (1) (5)	†23677·16 (3) (7)	22431·48 (2) (4)
	T		L.P.+	(He=0, Z=0)
Q′ (1)	26261·75 (*p*) (*p*)	24947·30 (*qd*) (?)	†23677·16 (3) (7)	22427·41 (*p*) (?)
	T	T $_4\delta_4$ Q (3) II_D *a*Q(7)	L.P.+	T
Q (2)	26275·27 (*rw*) (*r*)	24964·30 (*q*) (?)	23692·92 (0) (0)	22451·88 (1)* (3)
	D I (*r*)	T	T $_0\delta_1$ Q (2)	
Q′ (2)		24958·39 (1) (?)	23684·22 (*r*) (?)	22443·52 (*q*) (?)
		L.P.++180P (4)	D	D
Q (3)	26287·22 (2) (5)	24987·63 (0)* (2)	23722·26 (*p*) (0)	22489·69 (3) (5)
	L.P.++?20R (6)	$_4\delta_4$ Q (1)	T	Z, S
Q′ (3)		24976·78 (1) (2)	23717·59 (*r*) (?)	22481·81 (*q*) (?)
		L.P.++ S	D	T
Q (4)	26307·76 (*p*) (*p*)	25018·64 (*r*) (?)	23773·77 (5) (8)	22559·76 (0) (1)
	T	D	II_c *e* Q (5)	He++
Q′ (4)	26301·05 (*rd*) (?)	‡25013·51 (*p*) (?)	23768·56 (1) (?)	22555·95 (0) (?)
	D	T II_D *c*Q (1)	L.P.++	? 20P (4)
Q (5)	26346·71 (0) (1)	25074·73 (1) (3)	23847·79 (*q*) (*p*)	22657·81 (4) (8)
			T	He++, Z, S, $^1A_1$Q′(3)
Q′ (5)			23842·00 (*r*) (?)	22651·60 (1) (?)
			D	He++

* These lines are mixed with others on the f.f. discharge plate. Where these are in the same square the two lines are unresolved.

† Unresolved doublets (M. and B.).

‡ λ 3996·71 not completely resolved from λ 3997·18 (2) where measured by Tanaka.

Table I—(continued).

	$^3A_0$.	$^3A_1$.	$^3A_2$.	$^3A_3$.
Q (1)		27247·81 (*r*) (0)	25971·06 (0) (3)	24723·65 (1) (5)
		T	T $_3\beta_1$ Q (2)	II_D *b*Q (1)
Q′ (1)			25966·81 (*q*) (?)	24718·03 (*p*)
			T II_E *a*Q (5)	T $_0r_0$ Q (5)
Q (2)		27258·28 (*r*) (*r*)	25992·25 (*p*) §	24750·09 (*r*) (?)
		T	T	T
Q′ (2)			25981·65 (*rw*) (*rd*)	24744·21 (*rd*)
			D I (*r*) II_κ*b*Q (4)	T II_D *a*R (3) *b*R(1)
Q (3)		27283·57 (0) (1)	26023·91 (*r*) (*r*)	24790·03 (*r*) (?)
			D	T
Q′ (3)			26017·88 (*r*) (?)	
			T II_E*a*R (4)$_E$ *c*Q (3)	
Q (4)		27323·97 (*r*) (*p*)	26077·99 (*r*) (?)	24867·08 (*r*) (*r*)
		D	D	D
Q (4)			26072·76 (*r*) (*a*)	24867·08 (*r*) (*r*)
			D I (*r*) $_3\epsilon_3$ Q (3)	D
Q (5)		27379·26 (*rd*) (?)	26148·78 (*q*) (*r*)	24958·39 (1)* (5)
		D	T $_0\delta_0$ Q (4)	180P (4) $_2A_1$Q′(2)
Q′ (5)				24955·77 (1)* (5)
				$_2A_1$Q(1)

* These lines are mixed with others on the f.f. discharge plate. Where these are in the same square the two lines are unresolved.

§ λ 3846·21 (*p*) (T) mixed with λ 3845·53 (*qd*) (T) is enhancedto intensity (2) in f.f. discharge.

The initial vibration term differences are given in Table II.

Table II.—Initial Vibration Term Differences.

$m \rightarrow$	1.	2.	3.	4.	5.
${}_1A_0 - {}_0A_0Q\,(m)$	2359·88	2359·10	2352·92	2346·45	2342·98
${}_1A_0 - {}_0A_0Q'(m)$	2359·41	2355·97*	2352·11	2345·74	2339·38*
${}_1A_1 - {}_0A_1Q(m)$	2360·04	2360·26	2351·69	2347·65	2338·33*
${}_1A_1 - {}_0A_1Q'(m)$	2359·83	2362·58*	2348·25*	2346·41	2341·58
${}_1A_2 - {}_0A_2Q(m)$	2359·92	2358·40	2352·84	2346·39	2342·90
${}_1A_2 - {}_0A_2Q'(m)$	2359·87	2357·75*	2354·98	2346·31	2339·50*
Weighted means	2359·83	2359·10	2352·21	2346·41	2342·94 or 2339·49
Differences		0·73	6·89	5·80	3·47 or 6·92 ?
${}_2A_0 - {}_1A_0Q(m)$	2328·06*	2330·63	2320·54*	2305·88	2295·94
${}_2A_0 - {}_1A_0Q'(m)$	2331·84			2304·58	
${}_2A_1 - {}_1A_1Q(m)$	2335·02	2327·11	2321·14*	2301·90*	2297·29
${}_2A_1 - {}_1A_1Q'(m)$	2330·69*	2326·59	2318·97	2301·21*	
${}_2A_2 - {}_1A_2Q(m)$	2335·18	2327·27	2318·06	2305·42	2295·51
${}_2A_2 - {}_1A_2Q'(m)$	2334·77	2324·60*	2316·37*	2305·33	2295·85
${}_2A_3 - {}_1A_3Q(m)$	2333·73	2329·13	2316·68*	2306·76	2296·61
${}_2A_3 - {}_1A_3Q'(m)$	2333·82	2325·30	2313·04*	2304·59	2295·16
Weighted means	2335·10	2327·19	2318·06	2305·55	2295·81
Differences		7·91	9·13	12·51	9·74
${}_3A_1 - {}_2A_1Q(m)$	2292·04	2293·98*	2295·94*	2305·33	2304·43*
${}_3A_2 - {}_2A_2Q(m)$	2291·72	2299·33*	2301·65	2304·22	2301·00
${}_3A_2 - {}_2A_2Q'(m)$	2291·65	2297·43	2300·29	2304·20	
${}_3A_3 - {}_2A_3Q(m)$	2292·17	2298·21	2300·34	2307·32*	2300·58
${}_3A_3 - {}_2A_3Q'(m)$	2290·62	2300·69			2304·17*
Weighted means	2291·98	2297·43	2300·73	2304·21	2301·–
Differences		−5·45	−3·30	−3·48	+3·2–

The figures marked with an asterisk in Table II depend on one or more lines which are abnormally intense or otherwise irregular, and they have been left out of account in deciding on the probable correct values. Many of the cases which have an error of about 2 or 3 from the mean are probably due to the combination of the individual lines of one band with the two constituents unresolved in another. On the whole the agreement seems as good as is to be expected when the weakness and high wave-numbers of many of the lines is taken into consideration.

These initial vibration term differences are very peculiar. The $1 \rightarrow 0$ and $2 \rightarrow 1$ terms start by diminishing in a fairly normal way as "m" increases, but when "m" reaches about 3 or 4 the rate of diminution with increasing "m" begins to fall off. The $3 \rightarrow 2$ terms start by increasing and then diminish

as "m" increases. This is in marked contrast to the quite normal behaviour in this respect of the initial vibration differences of $2\sigma - m\pi$ described in "Structure," Part V. It is probably connected with the peculiar behaviour of the differences of the Q branches of these bands.

The final vibration term differences are given in Table III.

These figures show that the variation of the vibrational term differences with increasing rotational quantum number are, in contrast to the initial states, of almost normal type for the final states, the second differences being almost constant. Even here, however, there is an indication of a rise and subsequent fall of the second difference as "m" increases. As a first approximation I shall disregard this irregularity and analyse the data by the method used in "Structure," Part V (p. 408). For the $1 \rightarrow 0$ final vibration term differences the successive second differences 3·51, 4·44, 4·54, 5·7 and 2·9 have a mean value 4·22, for the $2 \rightarrow 1$ terms the second differences 3·18, 6·79 and 3·99 have a mean value 4·65, and for the $3 \rightarrow 2$ terms the second differences 4·22, 9·71 and 4·59 have a mean value 6·17. I take half the mean of the first two, which are in close agreement, viz., 2·22 as the value of $3/2\omega_0 u^2(1 + 2b)$. Here ω_0 is the vibration frequency at infinitesimal amplitude, $u = h/4\pi^2 J_0 \omega_0 c'$, J_0 the moment of inertia of the excited molecule in the rotationless vibrationless state, "c'" the velocity of light and "b" the constant in the expansion.

$$U = -a\left(\frac{1}{\alpha} + \frac{1}{\rho} - \frac{1}{2\rho^2} + b\varepsilon_1^3 + c\varepsilon_1^4 + \ldots\right)$$

(a, α, b, c constants, $\rho = r/r_0$, $\varepsilon = r/r_0 - 1$, r being the instantaneous and r_0 the equilibrium distance between the nuclei).

To determine ω_0 we have, from the 2, 1 and 0 vibration states Q (1) data, the relations $1/n\,(E_2 - E_1) = \omega_0\,(1 - 3x + \ldots) = 1276{\cdot}42$ and $1/n\,(E_1 - E_0) = \omega_0\,(1 - x + \ldots) = 1313{\cdot}26$. These give $\omega_0 x = 18{\cdot}42$, and, after adding the small correction for the effect of the rotational on the vibrational energy as in "Structure," Part V, $\omega_0 = 1336{\cdot}68$. The data under discussion are collected in Table IV.

The numbers in the third and fifth columns give only two equations to determine the three unknown quantities u, b, and c. If we assume $c = 0$, we get $u = 4{\cdot}17 \times 10^{-2}$; and if we assume $b = 0$, we get $u = 3{\cdot}515 \times 10^{-2}$ (and $c = -0{\cdot}738$). These values differ by about 20 per cent. The values in the table are those got by assuming $c = 0$. The corresponding values found in Part V for the 2σ state are given in the last row of the table. The closeness of the values of $u\omega_0$, J_0 and r_0 found for the present 2 state, and the 2σ state is not warranted

Table III.—Final Vibration Term Differences.

$m \rightarrow$	1.	2.	3.	4.	5.	6.	7.
${}_0A_0-{}_0A_1Q(m)$	1313·10	1308·61	1298·96	1286·34	1268·68		
${}_0A_0-{}_0A_1Q'(m)$	1313·72	1314·23*	1296·97	1284·84*	1266·66		
${}_1A_0-{}_1A_1Q(m)$	1312·94	1307·45	1300·19	1285·14	1273·33*	1244·95	1218·96
${}_1A_0-{}_1A_1Q'(m)$	1313·30	1307·62	1300 83	1284·17	1264·46*		1218·46
${}_2A_0-{}_2A_1Q(m)$	1305·98*	1310·97	1299·59	1286·16	1269·01		
${}_2A_0-{}_2A_1Q'(m)$	1314·45	—	—	1287·54			
Weighted means....	1313·26	1308·04	1299·31	1286·14	1268·43	1245·–	1218·71
1st differences.....		5·22	8·73	13·17	17·71	23·43	26·3–
2nd ,, ...		3·51	4·44	4·54	5·7	2·9	
${}_0A_1-{}_0A_2Q(m)$	1276·37	1269·68	1263·44	1247·13	1229·73		
${}_0A_1-{}_0A_2Q'(m)$	1276·26	1267·35*	1263·32	1248·97	1229·21		
${}_1A_1-{}_1A_2Q(m)$	1276·49	1271·54	1262·29	1248·39	1225·16*		
${}_1A_1-{}_1A_2Q'(m)$	1276·22	1272·18	1256·59*	1249·07	1231·29*		
${}_2A_1-{}_2A_2Q(m)$	1276·43	1271·38	1265·37*	1244·87*	1226·94*		
${}_2A_1-{}_2A_2Q'(m)$	1270·14*	1274·17*	1259·19*	1244·95*			
${}_3A_1-{}_3A_2Q(m)$	1276·75	1266·03*	1259·66*	1245·98	1230·47		
Weighted means	1276·42	1271·46	1263·32	1248·39	1229·47		
1st differences.....		4·96	8·14	14·93	18·92		
2nd ,, ...		3·18	6·79	3·99			
${}_1A_2-{}_1A_3Q(m)$	1246·51	1242·90	1231·19	1215·35	1191·08		
${}_1A_2-{}_1A_3Q'(m)$	1246·80	1241·40	1232·45	1211·87	1189·71		
${}_2A_2-{}_2A_3Q(m)$.....	1247·86	1241·04	1232·57	1214·01	1189·98		
${}_2A_2-{}_2A_3Q'(m)$	1247·75	1241·70	1235·78*	1212·61	1190·40		
${}_3A_2-{}_3A_3Q(m)$	1247·41	1242·16	1233·88	1210·91*	1190·40		
${}_3A_2-{}_3A_3Q'(m)$	1248·78	1237·44					
Weighted means....	1247·67	1241·70	1232·51	1213·61	1190·12		
1st differences.....		5·97	9·19	18·90	23·49		
2nd ,, ...		4·22	9·71	4·59			
${}_0B_0-{}_0B_1Q(m)$	1312·17	1310·80	1298·79	1290·32	1275·09*	1249·67	1222·31*
${}_0B_0-{}_0B_1Q'(m)$		1306·14	1297·38	1284·04	1266·96	1241·68	1216·06
${}_1B_0-{}_1B_1Q(m)$	1311·95		1312·14*				
${}_1B_0-{}_1B_1Q'(m)$	1314·86*						
Weighted means	1312·06	1308·47	1298·09	1287·19	1269·–	1245·68	1218·–
${}_0B_1-{}_0B_2Q(m)$	1276·53	1267·32*	1262·05	1249·55	1231·92		
${}_0B_1-{}_0B_2Q'(m)$	1278·03	1271·82	1264·14				
${}_1B_1-{}_1B_2Q(m)$	1277·01		1266·92*				
${}_1B_1-{}_1B_2Q'(m)$	1270·34*						
Weighted means	1277·19	1271·82	1263·10	1249·55	1231·92		
${}_0B_2-{}_0B_3Q(m)$.....	1248·03	1244·58	1236·16	1213·15	1188·5–		
${}_0B_2-{}_0B_3Q'(m)$	1246·69	1239·20	1228·12		1196·8–		
${}_1B_2-{}_1B_3Q(m)$	1253·48*		1228·76				
${}_1B_2-{}_1B_3Q'(m)$	1246·55						
Weighted means	1247·09	1241·9–	1232·14	1213·15	1190·–		

Table IV.

	ω_0.	$\omega_0 x$.	$\frac{2}{3}x = u \times 1+5b+c-\frac{5}{2}b^2$.	$\frac{3}{2}\omega_0 u^2 (1+2b)$.	$u\sqrt{1+2b}$.	b.	u.	$u\omega_0$.	J_0.	r_0.
2	1336·68	18·42	$0·920\times10^{-2}$	2·22	$3·315\times10^{-2}$	+2·146 or −0·146	$4·17\times10^{-2}$	55·7	$9·94\times10^{-41}$	$1·02\times10^{-8}$
3	2373·89	(12·36)	$(0·354\times10^{-2})$	(0·75)	$(1·453\times10^{-2})$	(2·148 or −0·148)	$(1·73\times10^{-2})$	(41)	—	—
2σ	2593·82	68·41	$1·757\times10^{-2}$	1·75	$2·122\times10^{-2}$	−0·0404	$2·216\times10^{-2}$	57·4	$9·65\times10^{-41}$	$1·086\times10^{-8}$

by the accuracy of the method and is likely to be fortuitous. In fact, as these bands are degraded towards the violet, it is probable that $u\omega_0$ will be appreciably lower and J_0 appreciably higher than for the 2σ states.

The values for the initial 3 state in the second row of Table IV have been got in a similar way, except that the values of $3/2\omega_0 u^2 (1 + 2b)$ are got from the vertical variation of the rotational second differences in Diagram I. This variation is small and the second differences are not constant and are subject to considerable errors, so that the method is very rough. The figure 41 for $u\omega_0$ should be about 13 units higher than the value for the final 2 state to agree with the second difference of the $0 \rightarrow 0$ lines. As both the rotational and vibrational second differences vary in a way which is not accounted for theoretically, it is doubtful how far the method is to be relied on, and the figures are only put in as indicating that the properties of the lines are not incompatible with the correct value of the moment of inertia of the molecule in the 3 state. It is possible that the puzzling features of the initial states might be due to something wrong with the arrangement of some of the lines, but this does not seem likely.

The distances between the doublets are shown in Table V.

The AQ (1) doublets behave quite regularly and show a steady increase in the doublet distance with increase in both initial and final vibration number. With the values of $m > 1$ there are so many interferences with these weak lines that it is difficult to be certain what these lines are doing, but they look to have an increasing separation as "m" increases. It is pretty certain that the existence of some staggering must be admitted if these doublets are real.

Apart from irregularities, all the Q branches of the A bands are of the type in which the second difference is equal to the first first difference, *i.e.*, the successive first differences are in the ratio 1 : 2 : 3 : 4 : 5, etc. This makes $\rho = +\frac{1}{2}$ in the formula $\nu = \nu_0 + (B - b)(m - \rho)^2$. In the $2\sigma - m\pi$ bands this ratio is 2 : 3 : 4 : 5, etc., requiring $\rho = -\frac{1}{2}$.

The B *bands.*

In addition to the A bands there is another system further out in the violet and extending into the ultra-violet which has the same final states as the A bands. It is less strongly developed than the A bands, but the leading lines are strongly enhanced in the f.f. discharge and there is an indication of doublets in its structure, though a number of them appear to have closed up. However, it is doubtful if the resolution which has been brought to bear on this part of the spectrum would separate these doublets in many cases. The first differences of the successive lines are approximately in the ratio 1 : 2 : 3, etc., so that

Table V.

$m \rightarrow$	1.	2.	3.	4.	5.	6.	7.
${}_0A_0 - {}_0A'_0$	3·31	2·19*	7·23*	4·70*	5·27		
${}_0A_1 - {}_0A'_1$	3·93	7·71*	5·24	3·20*	3·25		
${}_0A_2 - {}_0A'_2$	3·82	5·38	5·12	5·04*	2·73		
${}_1A_0 - {}_1A'_0$	3·78	5·22	8·04	5·41	8·87	4·77	6·20
${}_1A_1 - {}_1A'_1$	4·14	5·39*	8·68*	4·44	0·00*		5·70
${}_1A_2 - {}_1A'_2$	3·87	6·03*	2·98	5·12	6·13		
${}_1A_3 - {}_1A'_3$	4·16	4·53*	4·24*	1·64	4·76		
${}_2A_1 - {}_2A'_1$	8·47*	5·91*	10·85*	5·13*			
${}_2A_2 - {}_2A'_2$	0·00*	8·70	4·67	5·21	5·79*		
${}_2A_3 - {}_2A'_3$	4·07	8·36	7·88	3·81	6·21		
${}_3A_2 - {}_3A'_2$	4·25	10·60*	6·13	5·23			
${}_3A_3 - {}_3A'_3$	5·62	5·88			2·62*		
${}_0B_0 - {}_0B'_0$		10·50*	7·36	6·28	8·13	7·99*	11·19
${}_0B_1 - {}_0B'_1$	5·07	5·52	5·95	0·00*	0·00*	0·00*	4·94
${}_0B_2 - {}_0B'_2$	6·57*	9·84*	8·04*	0·00*	0·00*		
${}_0B_3 - {}_0B'_3$	5·23*	4·46	0·00*	0·00*	8·28*		
${}_1B_0 - {}_1B'_0$	11·19						
${}_1B_1 - {}_1B'_1$	13·60*						
${}_1B_2 - {}_1B'_2$	6·93						

The asterisks indicate that one or more abnormal lines is involved.

$\rho = +\frac{1}{2}$, as for the A bands. The only bands which are well developed are those with the initial vibration state 0, which I denote by ${}_0B_0$, ${}_0B_1$, ${}_0B_2$ and ${}_0B_3$. The bands with initial state 1, viz., ${}_1B_0$, ${}_1B_1$, ${}_1B_2$ and ${}_1B_3$, are each represented only by 2 lines, those for which $m = 1$ and $m = 3$. In the A bands and also in the ${}_0B_x$ bands the two strongest lines are, in general, those for which $m = 1$ and $m = 3$. The lines with their properties are set out in Table VI, in the same way as the lines of the A bands are treated in Table I.

Table VI.

	$_0B_0$.	$_0B_1$.	$_0B_2$.	$_0B_3$.
Q (1)	27133·88 (2) (7)	25821·71 (2) (7)	24545·18 (1) (2)	23297·15 (*q*) (?)
	L.P.+? $_0B_0$ Q (2)	L.P.++, He0, S		T $_2\delta_3$ Q (2)
Q′ (1)		25816·64 (1) (5)	24538·61 (1) (5)	23291·92 (0) (0)
		L.P.++, S	183P (3)	T
Q (2)	27143·38 (*r*) (?)	25832·58 (0) (0)	24565·26 (4) (8)	23320·68 (*r*) (?)
	DI (*r*) $_0B_0$ Q′ (3)		L.P.+	D
Q′ (2) ?	27133·88 (2) (7)	25827·04 (0) (?)	24555·42 (1) (4)	23316·22 (*p*) (*r*)
	L.P.+ $_0B_0$ Q (1)			T
Q (3)	27150·74 (1) (4)	25851·95 (1) (3)	24589·90 (1) (4)	23353·74 (*p*) (*p*)
		He0, S		T
Q′ (3)	27143·38 (*r*) (?)	25846·00 (*p*) (?)	24581·86 (4) (8)	23353·74 (*p*) (*p*)
	DI (*r*) $_0B_0$ Q (2)	T	L.P.+	T
Q (4)	27168·16 (*rd*) (*a*)	25877·84 (2) (5)	24628·29 (0) (1)	23415·14 (*qd*) (*r*)
	D	$_1\delta_1$ Q (1)		D
Q′ (4)	27161·88 (*q*) (*p*)	25877·84 (2) (5)	24628·29 (0) (1)	23415·14 (*qd*) (*r*)
	T	$_1\delta_1$ Q (1)		D
Q (5)	27189·21 (*r*) (*p*)	25914·12 (0) (0)	24682·2– [- (?)	23493·70 (*q*) (*r*)
	T		T	T $_1\delta_2$ Q (1)
Q′ (5)	27181·08 (*r*) (*p*)	25914·12 (0) (0)	24682·2– [– (?)	23485·42 (*q*) (*r*)
	T		T	T $_1\delta_2$ Q (2)
Q (6)	27207·25 (1) (7)	25957·58 (0) (3)		
	L.P.+	S		
Q′ (6)	27199·26 (*rd*) (*rd*)	25957·58 (0) (3)		
	DI (*r*)	S		
Q (7)	27233·56 (*rd*) (*r*)	26011·25 (*rd*) (*r*)		
	DI (*r*)	D		
Q′ (7)	27222·37 (*r*) (*r*)	26006·31 (*r*) (?)		
	T	D		

Table VI—(continued).

	$_1B_0$.	$_1B_1$.	$_1B_2$.	$_1B_3$.
Q (1)	29447·36 (1) (3)	28135·41 (*r*) (*a*)	26858·40 (0) (1)	25604·92 (*rd*) (?)
		T	L.P.+ 177P (4)	D
Q′ (1)	29436·17 (0) (?)	? 28121·81 (0) (*a*)	26851·47 (*r*) (*r*)*	25604·92 (*rd*) (?)
		L.P.+, 188P (3)	DI (*r*)	D
Q (2)				
Q′ (2)				
Q (3)	29472·27 (*rw*) (*rd*)	28160·13 (*rd*) (*a*)	26893·21 (*r*) (*p*)	25664·45 (*rw*)
	DI (*r*)	D	T	DI (*rw*)
Q′ (3)			26893·21 (*r*) (*p*)	
			T	

* λ 3723·14 (*r*) not completely resolved from λ 3723·68 (*r*).

The final vibration term differences of the B bands are given in Table III, and it will be seen that they agree satisfactorily with those for the A bands. The initial vibration differences are given in Table VI.

Table VI.

$m \downarrow$	$_1B_0 - {}_0B_0$	$_1B'_0 - {}_0B'_0$	$_1B_1 - {}_0B_1$	$_1B'_1 - {}_0B'_1$	$_1B_2 - {}_0B_2$	$_1B'_2 - {}_0B'_2$	$_1B_3 - {}_0B_3$	$_1B'_3 - {}_0B'_3$	Mean.
1	2313·48	—	2313·70	2305·17*	2313·22	2312·86	2307·77*	2313·00	2313·25
2									
3	2321·53*	—	2308·18	—	2307·33	—	2310·71	—	2307·76

and the doublet distances in Table VII.

To investigate the bands further I made the tentative assumption that the A bands correspond to the $3 \rightarrow 2$ and the B bands (as tabulated) to the $4 \rightarrow 2$ electron jumps. Then the magnitudes require that the $5 \rightarrow 2$ jump should give a C band with a nucleus very close to the line 29678·61 = 3368·47 (0) (0) L.P. +. This line occurs in f.f. discharge, but apparently is not enhanced. If this is $_0C_0$ Q (1), the $_0C_1$ Q (1) line should be at 28365·35 and does not occur in the tables. (It is coincident with a Hg line λ 3524·27 (2) — Diffuse 1P — 7D.) The line $_0C_2$ Q (1) could be 27090·95 (0) (1) = λ 3690·22 and the $_0C_3$ Q (1) line could be 25846·00 (*p*) (?) . T = λ 3867·98. There are no other lines,

Table VII.

$m\rightarrow$	1.	2.	3.	4.	5.	6.	7.
${}_0B_0-{}_0B'_0$		10·50*	7·36	6·28	8·13	7·99*	11·19
${}_0B_1-{}_0B'_1$	5·07	5·52	5·95	0·00	0·00	0·00	4·94
${}_0B_2-{}_0B'_2$	6·57*	9·84*	8·04*	0·00	0·00		
${}_0B_3-{}_0B'_3$	5·23*	4·46	0·00	0·00	8·28*		
${}_1B_0-{}_1B'_0$	11·19						
${}_1B_1-{}_1B'_1$	13·60*						
${}_1B_2-{}_1B'_2$	6·93						

but this is not surprising, as the electronic energy is getting very close to the ionisation potential value and these states must be very unstable.

Dieke and Hopfield's data on the states in the ultra-violet bands have been calculated on the formulæ of the wave-mechanics, so that in order to compare with their results I have calculated the null frequencies ν_0 of the A, B, and C band systems with the formulæ of the wave-mechanics. Strictly, there are only enough data to do this properly for the A bands, so for the B bands I assumed that the differences of the initial vibration frequencies to be the same as for the A bands, and I assumed values for the C bands extrapolated from the data for the A and B bands. These assumptions do not exert any profound influence on the results of the calculations. The null values were found to be: for A, $\nu_0 = 21050{\cdot}65$; for B, $\nu_0 = 26639{\cdot}49$; and for C, $\nu_0 = 29212{\cdot}98$. I then calculated a Hicks formula to fit these values of ν_0 and found it to be

$$\nu_0\,(m) = 33727{\cdot}12 - \frac{109678{\cdot}3}{\left(m - 0{\cdot}08926 + \dfrac{0{\cdot}09213}{m}\right)^2} \qquad m = 3,\ 4,\ 5.$$

The terms from this formula are $2S = 33727{\cdot}12$, $2P = (28643{\cdot}63)$, $3P = 12676{\cdot}47$, $4P = 7087{\cdot}66$, and $5P = 4514{\cdot}14$. These are similar to, but larger than, the corresponding values for the principal series of helium *singlets* which are $2S = 32032{\cdot}51$, $2P = 27175{\cdot}17$, $3P = 12100{\cdot}56$, $4P = 6817{\cdot}21$, $5P = 4367{\cdot}45$.

If this arrangement of the bands were correct, it would form a very interesting counterpart to Part V, in which the conclusion was reached that the band systems there dealt with had a main electronic structure closely resembling that of the principal series of helium *doublets*. It seems very doubtful, however, if this

arrangement is correct. The final vibrational differences of the Q (1) lines of the present bands are identical with the corresponding terms for the initial states of the Lyman bands (Dieke and Hopfield's B states) to the accuracy of the ultra-violet measurements. The ultra-violet data for lines other than Q (1) are not available at present, so that it is not certain the states are the same. However, the identity of the final states with Dieke and Hopfield's B states is confirmed by the values got for the heats of dissociation. If the states are the same, the present final electronic term 33727·12 should be equal to the difference between the ionisation potential of $H_2 = 128800$ and Dieke and Hopfield's electronic term $B - A = 91562$. This difference is 37238 and is much too large. I shall reconsider the matter in the next section.

The Heats of Dissociation, etc.

An estimate, possibly rough, of the heat of dissociation of the final 2 state of these bands can be got by the method which depends on the variation of the vibrational term differences with increasing vibration quantum number (see Part V, p. 415). Using the Q (1) lines the $1 \rightarrow 0$, $2 \rightarrow 1$ and $3 \rightarrow 2$ terms are 1313·26, 1276·42 and 1247·67 respectively. The successive differences are 36·84 and 28·75, and seem to be varying rapidly, but if the difference were to keep steady at 28·75, there would be $1247 \cdot 64/28 \cdot 75 = 43 \cdot 4$ additional states, making 46·4 altogether. At any rate, the number does not seem to be likely to be less than this, as the differences are diminishing with increasing n''. Taking the mean energy as half the energy of the first state, we get for the heat of dissociation in equivalent wave-numbers $1336 \cdot 68/2 \times 46 \cdot 4 \simeq 31{,}000$.

The data for the 2π state* treated in the same way give $2412 \cdot 18/2 \times 17 \cdot 62 \simeq 21300$. A check on this can be got by extrapolating from the values for the 3π, 4π, 5π and 6π states which are given in the papers cited.* One method gives about 17000 and the other about 24000. The mean of these is 20500, which confirms the value found directly. The final states of the present bands are evidently much more stable than the 2π states and also than any of the 2σ or $m\pi$ states described in Part V. I have shown in a preceding note† that there is some evidence which supports the view that the 2π states are the same as the C states of Dieke and Hopfield.‡ If the present final states are the same as Dieke and Hopfield's B states, we have a further check on these identifications. The difference between the electronic energy levels for the B and C

* O. W. Richardson, 'Roy. Soc. Proc.,' A, vol. 114, p. 648 (1927); also Part V, p. 415.

† 'Roy. Soc. Proc.,' vol. 114, p. 643 (1927).

‡ 'Z. f. Physik,' vol. 40, p. 299 (1926).

states should be equal to the difference of the heats of dissociation of molecules in the corresponding states, since all the 2 quantum states of the hydrogen atom have practically the same energy levels. If there is a difference, it will be within the limits set by the fine structure, if it has any, of the Lyman line-series line $2 \rightarrow 1$ (λ 1215·68). The differences of the electron levels for the ultra-violet bands given, in their notation, by Dieke and Hopfield are $B - A = 91562$ and $C - A = 99986$, so that $C - B = 8424$. The difference between the present estimates of the heats of dissociation of the molecules which are thought to be in the same two states is $31000 - 21300 = 9700$. These differences are the same to the degree of accuracy which can be claimed for the estimates of the heats of dissociation.

The heats of dissociation D_{M_y} of the initial states can be estimated from the electronic terms and the terms of Balmer's series by means of the equation $D_{M_y} = D_{M_x} + R_x A_y - R_x M_y$ (Part V, p. 414). The results are shown in Table VIII.

Table VIII.

	Initial states.					Final state.			
$m \rightarrow$	1.	2.	3.	4.	5.	1.	1.	2.	2.
$R_m A_\infty$	109768	27420	12187	6855	4387		109678		27420
$R_m A_{m+1}$	82258	15233	5332	2468	1340		82258		15233
$R_m M_\infty$	128800	28644	12677	7088	4514		128800	(37238)	33727
$R_m M_{m+1}$	100156	15967	5589	2574		(91562)	95073		
$R_m M_{m+1} - R_m A_{m+1}$	17898	734	257	106		(9304)	12815		
DM_m	34050	16152	15318	15212			34050	(24746)	21235
Equivalent, volts	4·21	2·00	1·90	1·87			4·21		2·63
Equivalent, cals.$=10\times^4$	9·72	4·62	4·38	4·32			9·72		6·06

In this table the value of the ionisation potential of the normal hydrogen molecule ($R_1 M_\infty$) used in Part V has been replaced by the experimental value 15·9 volts, since this agrees with the spectroscopic data of Dieke and Hopfield. The values are very similar to those found for the $m\pi$ states by this method as amended in 'Roy. Soc. Proc.,' A, vol. 114, p. 649 (1927). The same method applied to the term 2S of the Hicks formula gives for the heat of dissociation 21235 equivalent wave-numbers. It is very improbable that the estimate

(31000) got from the vibration terms can be in error by as much as this. The estimates in this and previous papers from the electron terms all come consistently low, and I wonder if the value of the heat of dissociation of normal H_2 may not be a little higher than the value found by Langmuir, or, alternatively, the spectroscopic value of the ionisation potential of H_2 might be a little under the experimental value 15·9 volts. Even so, the utmost that could be allowed in this way is probably about 5000 wave-numbers, which would still be insufficient to bridge the discrepancy. This brings us back to the difficulty that the calculated 2S term 33727 is too low to agree with the apparent identity of the final states with Dieke and Hopfield's B states. If their corresponding electronic term 37238 is used instead of 33727, the values in brackets in Table VIII are obtained. The value of the heat of dissociation 24746 is now just about as much below the value estimated from the vibration term data as these estimates from the electronic term data generally seem to run.

These difficulties can be avoided by the following re-arrangement of the bands, which, however, can only be regarded as provisional at this stage. I still take the strong A band system as a $3 \rightarrow 2$ electron transition, but I take out the ${}_0B_0$, ${}_0B_1$, ${}_0B_2$ and ${}_0B_3$ bands and the C fragment and assume the ${}_1B_0$, ${}_1B_1$, ${}_1B_2$ and ${}_1B_3$ lines to be respectively the ${}_0B_0$, ${}_0B_1$, ${}_0B_2$ and ${}_0B_3$ lines of the $4 \rightarrow 2$ electron transition bands. The electron levels now proposed are shown diagrammatically, but not to scale, in Diagram II. The three lowest levels are the A, B and C levels of Dieke and Hopfield, and the others give the surface of the molecule corresponding to the ionisation potential of normal H_2 and the levels for the new ${}_0B_0$ Q (1) (${}_1B_0$ Q (1) of Table VI) and the ${}_0A_0$ Q (1) lines. The higher levels are getting near the ionisation potential limit, and it is not surprising if they are not found.

In what follows the experimental value 128800 of the ionisation potential is not used, but a corrected value 129429 got by adding a value $2\pi = 29443$ to 99986, which is Dieke and Hopfield's value of C — A.

If the foregoing arrangement is correct, it seems to suggest some rather surprising consequences. If the final electronic term of the present bands is 37867, we have for the initial term 37867 — 21574 = 16293, and for the initial term of the B fragment 37867 — 29447 = 8420. These numbers are all very close to the $m\sigma$ terms of the helium doublets. The mσ and $m\pi$ terms of the helium *doublet* band system are collected in Table IX, together with the above terms and also the initial and final electronic terms of the bands in Part V.

DIAGRAM II.

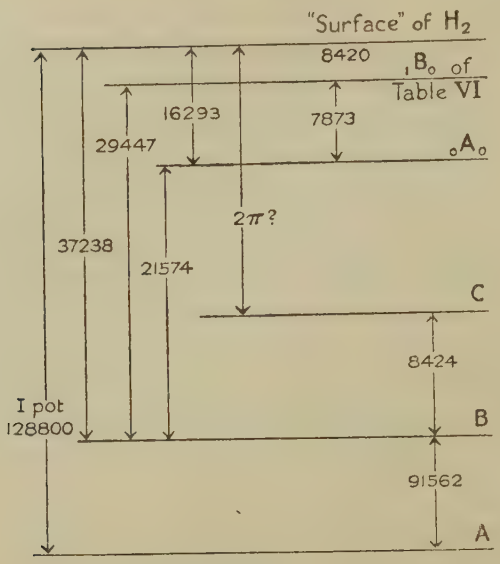


Table IX.—Electronic Terms.

	Final.	Initial.				
$m \rightarrow$.	2.	3.	4.	5.	6.	7.
Present H_2 bands	37867	16293	8420			
He $m\sigma$	38454·71	15073·87	8012·53			
H_2 bands in Part V	29502·82	12718·87	7067·07	4492·78	3106·80	2275·86
He $m\pi$	29222·85	12746·08	7093·59	4509·96	3117·66	2283·31

This table suggests that the present bands are of the type $2\sigma - m\sigma$ and the bands in Part V of the type $2\pi - m\pi$ or perhaps $2\pi_{1 \text{ or } 2} - m\pi_{2 \text{ or } 1}$, and not $2\sigma - m\pi$, as previously supposed.

It is possible that the $_0B_0$, $_0B_1$, $_0B_2$ and $_0B_3$ bands which have been taken out are a progression of vibration transitions of another band system with the same final state.

I hope to return to these questions as more evidence accumulates.

In conclusion I should like to thank Dr. D. B. Deodhar for his help in the examination of the spectrum plates.

The Magnetic Anisotropy of Crystalline Nitrates and Carbonates.

By K. S. Krishnan and Prof. C. V. Raman, F.R.S.

(Received March 27, 1927.)

1. *Introduction.*

Recent work in the field of magnetism emphasises the relation between the crystal structure and magnetic properties of solids. Such relations exist in all crystals whether ferromagnetic, paramagnetic or diamagnetic. We propose in this paper to discuss the explanation of the anisotropy of diamagnetic crystals, particularly some nitrates and carbonates with regard to which we have data from the measurements of Voigt and Kinoshita* and very recently of Rabi.† These substances are especially simple because, as is well-known from the work of Kossel, Bragg and others, they consist of charged ions held together by electrostatic forces, so that we can attribute the magnetic anisotropy of the crystal to that of the individual ions. Thus, for instance, in the case of sodium and potassium nitrates we may reasonably look for the explanation of observed magnetic anisotropy of the crystal in the structure of the nitrate ion, since presumably the metallic ions are more or less isotropic. It is significant in this connection that nitric acid solution has been found by Cotton and Mouton‡ to exhibit a measurable magnetic birefringence, thus indicating that the NO_3^- ion has a pronounced magnetic anisotropy even in the liquid state. It is possible to estimate this anisotropy in a purely optical way by combining the data for magnetic double-refraction with the measurements of the depolarisation of the light scattered in nitric acid. It is found that the magnetic anisotropy of the NO_3 group thus found, practically agrees with that necessary, as remarked above, to explain the magnetic properties of the crystalline nitrates.

2. *Relation between Structure and Magnetic Properties.*

The crystal structures of sodium nitrate and calcite have been investigated by X-ray methods and the two crystals are found to be homœomorphous. They have an axis of trigonal symmetry. The NO_3 or CO_3 groups are built up of three oxygen atoms at the corners of an equilateral triangle with the nitrogen or carbon atom at the centre, the plane of the triangle being perpendicular to

* 'Ann. d. Physik,' vol. 24, p. 492 (1907).
† 'Phys. Rev.,' vol. 29, p. 174 (Jan., 1927).
‡ 'Ann. Chim. et Phys.,' vol. 28, p. 209 (1913).

the trigonal axis. Potassium nitrate and aragonite, on the other hand, belong to the rhombic system; but here also the three oxygen atoms are arranged symmetrically round the nitrogen or carbon atoms in a plane perpendicular to the "c" axis.

In the following table are given the principal diamagnetic susceptibilities for these crystals :—

Compound.	Crystal system.	Axis along which susceptibility is measured.	Susceptibility per gm. molec. $\times -10^6$.	Average susceptibility per gm. molec. $\times -10^6$.
$NaNO_3$	trigonal	$\parallel$ trig. axis	29·5	25·9
		$\perp$ trig. axis	24·1	
$CaCO_3$ (calcite)	trigonal	$\parallel$ trig. axis	40·6	37·8
		$\perp$ trig. axis	36·4	
KNO_3	rhombic	$\parallel$ "c" axis	35·6	31·7,
		$\parallel$ "b" axis	29·7	
		$\parallel$ "a" axis	29·9	
$CaCO_3$ (aragonite)	rhombic	$\parallel$ "c" axis	44·4	40·8
		$\parallel$ "b" axis	38·7	
		$\parallel$ "a" axis	39·2	

Sodium nitrate and calcite are magnetically uniaxial owing to the possession of an axis of trigonal symmetry. For potassium nitrate and aragonite, however, the susceptibilities along the "b" and "a" axes are different. But the difference is very small, so that for all practical purposes we may treat them also as magnetically uniaxial crystals. In all the cases quoted in the table, the susceptibility along the axis, which we shall denote by χ, is numerically greater than the susceptibility, $\chi_\perp$, perpendicular to the axis. It is also significant that $\chi_\parallel - \chi_\perp$ is very nearly the same for the two nitrates, and not very different for the two forms of calcium carbonate. It suggests immediately that the nitrate and carbonate ions are, at least in a very large measure, responsible for the magnetic anisotropy of the respective crystals.

3. *Relation to Magnetic Birefringence.*

Before proceeding to connect the magnetic anisotropy of the NO_3^- ion in the crystals with its anisotropy as deduced from observations on magnetic double-refraction of nitric acid (liquid), we may point out that we would not be justified in assuming that the NO_3^- ion is necessarily identical in the two cases. For instance, it is well-known from the investigations of Oxley* that many dia-

* 'Phil. Trans.,' A, vol. 214, p. 109 (1914).

magnetic substances show an appreciable change of susceptibility on crystallisation, suggesting a re-adjustment of the electron-orbits as we pass from the liquid to the crystalline state. We cannot therefore expect any exact numerical agreement between the values of the anisotropy of the NO_3^- ion calculated from the two states.

The constant of magnetic birefringence (Cotton–Mouton constant) of nitric acid has been measured by Cotton and Mouton and $= 6{\cdot}3 \times 10^{-14}$ at about 16° C. for light of wave-length $\lambda = 5{\cdot}78 \times 10^{-5}$ cm. They, however, used acid of density 1·49, which corresponds to a concentration of about 90 per cent. Since the magnetic birefringence of water is negligibly small, we will not be far out if we estimate the value of the constant for 100 per cent. acid at

$$6{\cdot}3 \times 10^{-14} \times \frac{10}{9} = 7{\cdot}0 \times 10^{-14}.$$

In order to evaluate the magnetic anisotropy of the molecule from the Cotton-Mouton constant, we require to know the optical constants of the molecule. From the recent work of Bragg* on the birefringence of crystalline nitrates, we have very strong reasons to believe that HNO_3 molecule possesses an axis of optical symmetry, which is perpendicular to the plane of the NO_3^- ion; the moment induced in the molecule by unit field of the incident light-waves, acting along this axis, equal to C, say, being less than when it is acting perpendicular to it (= A). On actual calculation from Bragg's data, the ratio of these moments C/A comes out equal to 0·60. However, since the Cotton-Mouton constant refers to the liquid state it is only proper that we should use the value of the optical anisotropy calculated from measurements under the same conditions. Recently careful measurements have been made in the authors' laboratory by Mr. S. Venkateswaran† on the depolarisation of the light scattered by nitric acid of different concentrations. The value of the depolarisation factor extrapolated for 100 per cent. acid = 0·64. Calculating from this value we get

$$C/A = 0{\cdot}38 \text{ or } = 0{\cdot}50,$$

according to the two hypotheses that have been proposed for light-scattering. Even though the experimental data at present available are not sufficient to decide definitely which of the two hypotheses is correct, the data are more in accord with the hypothesis which gives the latter value of C/A, viz., 0·50.‡

* 'Roy. Soc. Proc.,' A, vol. 106, p. 346 (1924).

† 'Indian Journal of Physics,' vol. 1, p. 235 (1927).

‡ See K. S. Krishnan, 'Proc. Indian Ass. Sci.,' vol. 9, p. 251 (1926).

Now the magnetic anisotropy of HNO_3 molecule can be calculated from the Cotton-Mouton constant, C_m of liquid nitric acid, with the help of the relation

$$C_m = -\frac{(n_0^2-1)\,(n_0^2+2)}{30 n_0 \lambda k T} \cdot \frac{A-C}{2A+C} \cdot (C'-A'),$$

where n_0 is the refractive index of the liquid outside the field, k is Boltzmann's constant, C′ and A′ are the susceptibilities of the molecule along and perpendicular to the axis respectively. A being greater than C, it is evident from the above expression that in order to get a positive value for magnetic birefringence, as is actually observed, C′ should be numerically greater than A′, *i.e.*, the diamagnetic susceptibility of the NO_3^- ion along its axis should be numerically greater than for perpendicular directions—in conformity with the observations on crystals.

Actually substituting values for C_m and C/A in the expression, we obtain

$$C' - A' = -6{\cdot}7 \times 10^{-30}$$

$$\text{or } -8{\cdot}8 \times 10^{-30}$$

corresponding to the two values of C/A.

Now according to the assumptions we have made

$$\chi_{\parallel} = N\,C'$$

and

$$\chi_{\perp} = N\,A'$$

where N is the Avogadro number per gram molecule.

Therefore $\chi_{\parallel} - \chi_{\perp} = -4{\cdot}1 \times 10^{-6}$

or $-5{\cdot}3 \times 10^{-6}$,

which may be compared with the observed values

$-5{\cdot}4 \times 10^{-6}$ for $NaNO_3$

and $-5{\cdot}8 \times 10^{-6}$ for KNO_3.

It may also be mentioned incidentally that the absolute value of the susceptibility of NO_3^- ion is the same in the crystals and in nitric acid. If from the average susceptibilities of $NaNO_3$ and KNO_3 we deduct the contributions from the Na^+ and K^+ ions as given by Joos,* we get for the average susceptibility of NO_3^- ion respectively the values—

$$\chi = -(25{\cdot}9 - 6{\cdot}5) \times 10^{-6} = -19{\cdot}4 \times 10^{-6}$$

and

$$\chi = -(31{\cdot}7 - 14{\cdot}5) \times 10^{-6} = -17{\cdot}2 \times 10^{-6}$$

per gram molecule.

* G. Joos, 'Z. f. Physik,' vol. 32, p. 885 (1925).

The value calculated from Quincke's results for nitric acid of about 63 per cent. concentration

$$= -17 \times 10^{-6} \text{ per gm. molecule.}*$$

4. *Relation to the Structure of Ions.*

We have to look for the explanation of the observed magnetic anisotropy of the NO_3^- ion in its peculiar electronic structure. The three O-atoms are distributed symmetrically round the N-atom as centre, all of them lying in the same plane. Pauling† has recently suggested a dynamical model from entirely independent considerations, where he assumes six of the electrons to move in pairs in orbits connecting the central N-atom in turn with the three O-atoms, the other eighteen electrons being distributed about the three O-atoms. (Of course we exclude the K-electrons as contributing negligibly to the susceptibility.) If we now assume that the electron-orbits connecting the N- and O-atoms are in the plane of the atoms, and the other orbits are orientated more or less isotropically, the diamagnetic susceptibility perpendicular to the plane of NO_3^- ion will be numerically greater than for directions in the plane by an amount equal to the contribution from the six binuclear orbits. Taking the other electron-orbits to be all of equal size, it is found on calculation, from the assumption made above, that the susceptibility of a binuclear orbit should be much less than that of one of the other orbits. Too little is known at present regarding multinuclear electron-orbits to enable us to verify how far this conclusion is in accord with the general theory of diamagnetism. The results seem to suggest that these binuclear orbits, if elliptical, should have a large eccentricity.

However, the main conclusion that the anisotropy arises from the six binuclear orbits, seems to gain a further support from the fact that the CO_3^- ion, which has an essentially similar structure, exhibits a magnetic anisotropy (defined by $\chi_{\parallel} - \chi_{\perp}$) almost the same as for the NO_3^- ion. In this connection it need hardly be emphasised that further data on the magnetic susceptibilities of a number of crystals of this type, *e.g.*, nitrates, silicates, carbonates and borates, would be highly desirable and we have initiated experimental work in this direction.

* The susceptibility of nitric acid is assumed, consistently with our assumptions, to be entirely due to the NO_3 ion, the hydrogen atom, having lost its electron, contributing nothing to it.

† L. Pauling, 'J. Am. Chem. Soc.,' vol. 48, p. 1139 (1926).

5. *Summary.*

Crystals of sodium and potassium nitrates exhibit a marked diamagnetic anisotropy, the susceptibility perpendicular to the plane of the NO_3^- ion being greater than for directions in the plane ; the difference of susceptibility in the two directions is the same for the two crystals.

Attributing this anisotropy to that of the NO_3^- ion, it is found that its magnitude is exactly what we should expect from the known value of the magnetic birefringence (Cotton-Mouton effect) of nitric acid liquid.

An explanation is suggested on the basis of its electronic structure ; CO_3^- ion which has essentially the same structure, gives almost the same anisotropy.

A New Differential Dilatometer for the Determination of Volume Changes during Solidification.

By C. J. Smith, Ph.D., M.Sc., A.R.C.S., Metallurgical Research Laboratory, University of Sheffield.

(Communicated by C. H. Desch, F.R.S.—Received April 6, 1927.)

In recent years, the practical requirements of the metal industries have made it necessary to study the factors which govern the production of good castings. One of the most important of these factors is the change of volume which accompanies solidification. The experimental methods which have hitherto been used to determine this change have given discordant results, and it has seemed desirable to devise a new method, less liable to error. The new form of volumenometer which is the subject of this paper is intended to eliminate most of the errors inherent in the older methods. It has been applied to the measurement of the volume changes of two eutectic alloys, those of lead and tin and of tin and bismuth, the former of which contracts during solidification, whilst the latter shows a distinct expansion. The results indicate that the method is trustworthy.

Previous Methods of Measurement.

The older methods, which have been used for the experimental determination of the changes in volume, associated with the change in state of bodies, may be divided into the following classes :—

(*a*) The coefficients of expansion of the solid and liquid, over limited ranges

of temperature, are measured and the volume change occurring at the melting point is found by extrapolation. The coefficient of expansion of the solid is found either by direct measurement of the linear expansion or deduced from measurements derived from some hydrostatic method in which Archimedes' Principle is employed. The expansion of the liquid melt is inferred from observations on some dilatometric or hydrostatic method.

Objection has been raised to this general method, since the estimation of the change in volume which occurs at the melting point depends upon the unwarranted extrapolation of the experimental results, and it is only in the case of a pure metal or a eutectic that the change occurs at one definite temperature, namely, the melting point.

As an example of the discrepancies which occur in various expansion determinations (λ), we may cite the case of zinc (see Table I).

Table I.

Temperature or temperature range.	$\lambda \times 10^5$.	Authority.	Reference.
° C.			
40	2·92	Fizeau	'C.R.,' vol. 68, p. 1125 (1869).
10	2·97	Dorcez	'Phys. Rev.,' vol. 25, p. 88 (1897).
0–100	1·65	Grüneisen	'Ann. Phys.,' vol. 4 (1910).
20–100	3·65	Schulze	'Phys. Zeit.,' p. 403 (1921).
25–250	2·81	Smirnoff	'C.R.,' vol. 155, p. 352 (1912).
10–100	3·20	Friend and Vallance	'Jour. Inst. Met.' (1) (1924).

If we neglect the low result obtained by Grüneisen, the above figures are discordant amongst themselves to an extent of 30 per cent. Similar discordances occur in the figures relating to the liquid metal.

(*b*) Dilatometer methods have been employed by several experimentalists.* Such methods are very suitable for a study of the thermal expansion of metals of low melting point. The method is one in which the volume change is determined directly, and provided the metal is caused to freeze in the appropriate manner, such that cavities are avoided, the measurements are only subject to an error of 2 or 3 per cent. As the temperature at which the experiments are performed increases, it becomes very difficult to find a suitable liquid and

* Wiedmann, 'Ann. d. Phys. u. Chem. Beibl.,' vol. 20, p. 228 (1883); Toepler, *ibid.*, vol. 53, p. 343 (1894); Leduc, 'C. R.,' vol. 142, p. 46 (1906); Hackspill, 'C. R.,' vol. 152, p. 259 (1911); Bernini and Cantoni, 'Il Nuovo Cimento,' vol. 8, p. 241 (1914); Bornemann and Sauerwald, 'Z. f. Metallk.,' vol. 14, p. 145 (1922).

container. Most liquids of high boiling point show a variation of density at a fixed temperature which is an unknown function of the time for which that temperature has been maintained, and also depends upon the past thermal history of the liquid. At still higher temperatures, when ordinary glasses begin to deform, quartz vessels become permeable to organic liquids and salt mixtures attack the quartz. That the method yields consistent results at low temperatures is indicated by the recent work of Fleischmann, who has examined the volume changes occurring in the case of Rose's Metal. He used a dilatometer, the liquid being paraffin oil.*

(*c*) Methods, which cannot be classed as precision methods, consist in filling a mould with the metal under investigation and observing the variation with falling temperature of the distance between two pins embedded in the casting. It is difficult to know how to interpret these variations, but such results, whilst they do not measure the changes in volume occurring during solidification, have given information which foundrymen have found useful.

(*d*) Still more recent are the experiments of Endo,† who has elaborated a method which was suggested by Honda.‡ Essentially the method consists in finding the change in buoyancy of a metal which is suspended in an inactive liquid. These changes are observed by means of a thermobalance, which is fully described by the above workers, but the following account is rendered necessary if the ensuing criticism is to be appreciated :—

The metal is contained in a silica vessel and the whole is placed in a suitably chosen liquid maintained at an appropriate temperature. The vessel and its contents are suspended from one arm of the thermobalance, whilst suitable masses form a counterpoise. The results depend upon the measurement of the density of the liquid medium, and the density is not constant but varies with the time and temperature. In addition to this, the temperature of the metal is not equal to that of the liquid in which it is suspended, the difference amounting to several degrees. Furthermore, there exist in the bath itself temperature gradients which amount to a few degrees, and these cannot be eliminated by any method of stirring, for it is at once apparent that the equilibrium of the balance would be destroyed by any such process. In his paper Endo has not given any figures which are relevant to the above criticism, but his method has been employed in this department and the temperature difference has been found to be at least 10° C. across the bath. These variations manifest themselves in the

* ' Z. f. Physik,' vol. 41, p. 8 (1927).
† ' J. Inst. Metals,' vol. 30, p. 121 (1923).
‡ ' Sci. Rep. Tôhoku Univ.,' vol. 4, p. 97 (1915).

curves which are obtained when such methods are employed, *e.g.*, the curve in fig. 1 was obtained by some workers for the British Non-ferrous Metals Research

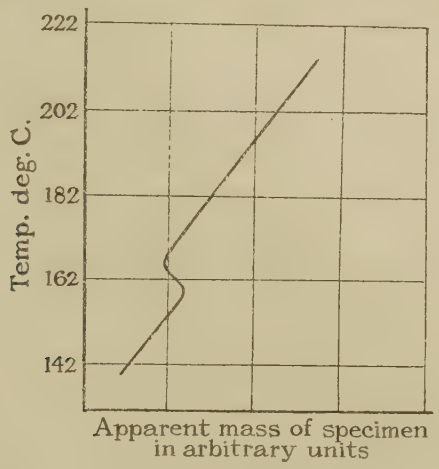


Fig. 1.

Association, and I am indebted to Dr. O. F. Hudson for permission to use this figure. The results would indicate that the melting point of the lead-tin eutectic is 162°, whereas it was known to be 181°–183° C.,* and that the volume changes occur over a definite temperature range. The lag is due to the fact that the thermocouple, used to measure the temperature, is outside the graphite vessel which contains the alloy, and the peculiar shape is probably due to effects associated with the latent heat of solidification. In fact, the indications from this method seem to be most unreliable in the region where accurate measurements are required.

The Theory of the Differential Dilatometer.

As a simple theoretical example of the method about to be described, let us consider two bulbs of volumes V and Ω, the former containing a solid of volume S. The two bulbs are connected together by means of a pressure gauge. The suffixes attached to the above symbols will indicate the appropriate volumes at temperatures corresponding to these suffixes. The two bulbs V and Ω can be maintained at any desired constant temperature by using a suitable thermostat. Let p and Π denote the pressures of the superincumbent gas in both bulbs. Then, since the mass of gas in each bulb is constant,

$$\frac{p_0 (V_0 - S_0)}{T_0} = \frac{p_T (V_T - S_T)}{T},$$

* Landolt and Bornstein Tables.

and $$\frac{\Pi_0\Omega_0}{T_0} = \frac{\Pi_T\Omega_T}{T},$$

whence

$$\frac{p_0(V_0 - S_0)}{\Pi_0\Omega_0} = \frac{p_T(V_T - S_T)}{\Pi_T\Omega_T}. \quad (1)$$

From this equation the volume S_T can be found, providing that the pressure inside the apparatus is known.

First Form of Dilatometer.

The first form of dilatometer which was constructed in the present research is shown in fig. 2. The apparatus was constructed entirely of soda glass, and

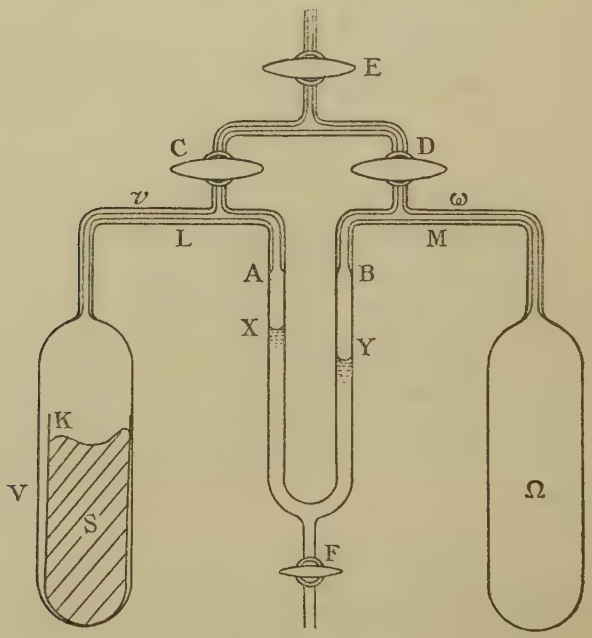


Fig. 2.

the metal was contained in a small glass vessel K. This container was necessary to guard against the cracking of the bulb V when solidification occurred ; as arranged here, only the container would be broken in such a case and the experiments could be completed. Although the figures show the apparatus to be in one plane, yet in practice this is undesirable. The capillary tubes which connected the bulbs V and Ω to the gauge were bent back at right angles at the points indicated by L and M. This enabled the bulbs to be in the same liquid bath and thus, after due precaution, to be at the same temperature.

In addition, that troublesome and variable factor, the temperature of the connecting tubes, is the same for each side of the apparatus, and, when this is so, the precision with which this temperature must be known is not necessarily high, for the differential nature of the volumenometer renders the apparatus almost self-compensating as far as the volumes of these connecting tubes are concerned. The stop-cocks C, D and E enable the apparatus to be exhausted completely and then filled with pure dry nitrogen. The pressure of the gas in the apparatus approximates to one atmosphere, and when the temperature is constant everywhere, the taps C and D are closed. The levels of the gauge liquid in AX and BY are now identical, and if steady temperature conditions have been really maintained at all points, then the levels at X and Y remain fixed. In these experiments the level did not change after periods of time never less than 1 hour.

Now let v and ω denote the volumes of the connecting tubes exposed to temperatures less than that of the thermostat, up to the points A and B; let l and λ be the volumes corresponding to AX and BY respectively. Let t denote the effective temperature of the connecting tube; strictly speaking, one requires the evaluation of $\Sigma \frac{dv}{t}$ in all this region, where dv is the element of volume of the tube at temperature t.

An application of the elementary kinetic theory then gives

$$\frac{p_0(V_0 - K_0 - S_0)}{\tau_0} + \frac{p_0 v_0}{t_0} + \frac{p_0 l_0}{\tau_0} = \frac{p(V - K - S)}{T} + \frac{pv}{t} + \frac{pl}{\tau}$$

and

$$\frac{\Pi_0 \Omega_0}{T_0} + \frac{\Pi_0 \omega_0}{t_0} + \frac{\Pi_0 \lambda_0}{\tau_0} = \frac{\Pi \Omega}{T} + \frac{\Pi \omega}{t} + \frac{\Pi \lambda}{\tau} \quad [\text{note } \Pi_0 = p_0].$$

We then have

$$\frac{p_0(V_0 - K_0 - S_0) + \frac{T_0}{t_0} v_0 + \frac{T_0}{\tau_0} l_0}{\Omega_0 + \frac{T_0}{t_0}\omega_0 + \frac{T_0}{\tau_0}\lambda_0} = \frac{p}{\Pi}\left[\frac{(V - K - S) + \frac{T}{t} v + \frac{T}{\tau} l}{\Omega + \frac{T}{t}\omega + \frac{T}{\tau}\lambda}\right], \quad (2)$$

from which equation S can be found providing that the pressure in the apparatus is known. Rather than insert a barometrical column with all its attendant difficulties and complications in the apparatus, it was considered better to deduce the pressure from the known volumes of the bulbs, etc., and the appropriate value of T, in order to find Π, and then to calculate p from this value for Π and the difference in pressure indicated by the height of the gauge XY. For this purpose it is essential to know the density of the oil in the gauge. The gauge

liquid was medicinal paraffin, this being chosen on account of its stability and low vapour pressure.

The Second Form of Differential Dilatometer.

During the course of this work the above apparatus was broken, and in the second form several additions were made, so that the calculation became more simple and dependent upon a smaller number of variables, owing to the difference in pressure between the two bulbs being made zero. This was brought about by inserting two tubes containing mercury (see fig. 3). These tubes

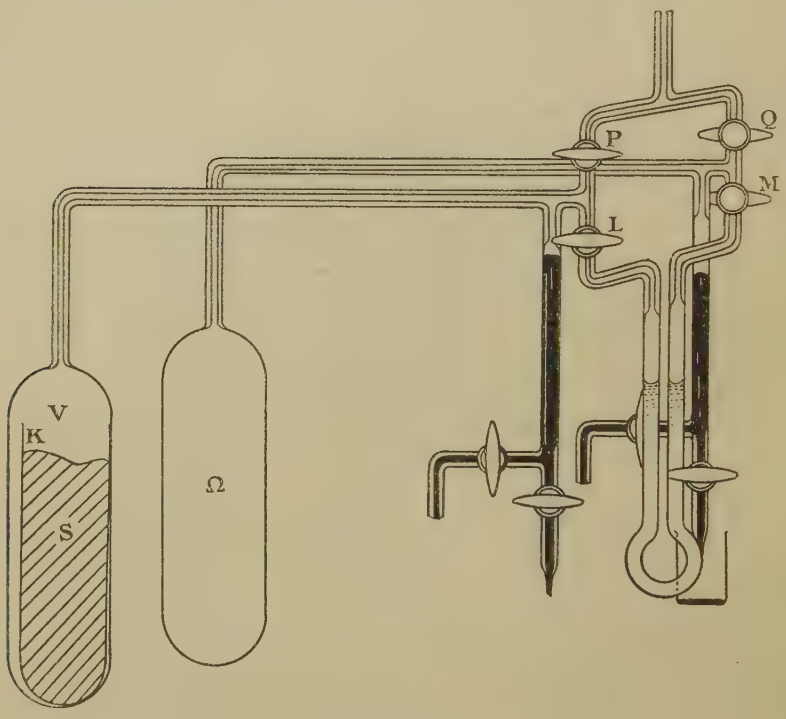


Fig. 3.

enable mercury to be abstracted from the appropriate side of the dilatometer until the above conditions are attained. It is at once apparent that the density of the oil in the gauge does not occur in the final expression from which S is calculated.

Let v and ω now be the volume of the connecting tubes and portions of the gauge not containing liquid.

Then, as before,

$$\frac{p_0(V_0 - K_0 - S_0)}{T_0} + \frac{p_0 v_0}{t_0} = \frac{p\,(V - K - S)}{T} + \frac{pv}{t},$$

and

$$\frac{p_0\,(\Omega_0)}{T_0} + \frac{p_0\omega_0}{t_0} = \frac{p\Omega}{T} + \frac{p\omega}{t} + \frac{p\delta}{t'},$$

where δ represents the increase in volume on the Ω side, due to the withdrawal of the mercury at temperature t' to render the pressure difference between the bulbs V and Ω zero. We then have

$$\frac{V_0 - K_0 - S_0 + \frac{T_0}{t_0} v_0}{\Omega_0 + \frac{T_0}{t_0}\omega_0} = \frac{V - K - S + \frac{T}{t} v}{\Omega + \frac{T}{t}(\omega) + \frac{T}{t'}\delta}. \qquad (3)$$

Even a cursory examination of this formula will reveal two advantages which this second form of the apparatus possesses over the earlier type. It is not necessary to measure the actual pressure inside the apparatus, and the temperature only occurs explicitly as a small correcting factor, due to the expansion of the glass, and due to the fact that the volumes of the connecting tubes must necessarily be finite. The temperature was measured by means of a platinum thermometer P, fig. 5.

During the course of this work a paper has appeared by Fr. Henglein* in which she has examined the volume changes associated with change in the temperature of some alkali halides. Fr. Henglein used a gas-filled dilatometer, but it was not a differential one. In her apparatus it was essential to measure the pressure by means of a column of mercury, and the difficulties which are attendant upon an accurate estimation of the lengths of such columns are well known. Furthermore, in her work it was necessary to know the temperature, for it occurs explicitly; the volume change associated with the glass has to be known more accurately, for there is no tendency for it to be compensated, as in the present work; lastly, the volumes of the connecting tubes do not tend to compensate one another as they do here.

Preparation of the Nitrogen.

The nitrogen which was requisite for the above experiments was prepared by a method due to Waran† and developed by the author.‡ The mode of

* 'Z. f. Phys. Chem.,' vol. 115, p. 91 (1925).

† H. P. Waran, 'Phil. Mag.,' vol. 41, p. 246 (1921).

‡ C. J. Smith, 'Proc. Phys. Soc.,' vol. 34, p. 155 (1922).

procedure is fully explained in the published papers, so that no further remarks are now necessary.

Preparation of the Metallic Eutectics.

The alloys which were used in the experiments were made from constituents which were of the highest attainable degree of purity. The tin was the so-called "Chempur" variety as supplied by Messrs. Capper and Pass. The bismuth was of a similar purity, whilst the lead was that which is generally known as "Assay" lead. In preparing the eutectics, about 400 gm. of alloy were generally made, the constituents being in the proportions as recorded in Landolt and Bornstein's Tables.

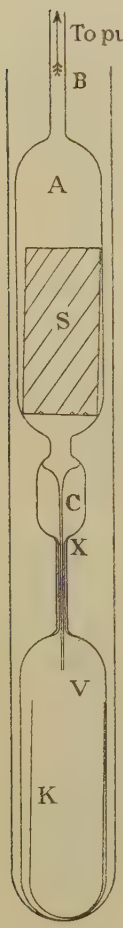


Fig. 4.

In each case the constituents were intimately mixed in a pyrex glass container and melted in a high-frequency induction vacuum furnace. The vacuum was certainly less than 10^{-3} mm., so that the alloys were obtained in as compact form as possible and free from cavities. Even after this treatment a slight scum always appeared on the surface of the metal; this was removed by turning a cylinder in a lathe. To introduce this specimen as a whole into the bulb V was impossible, and so a method was employed by which the process was accomplished *in vacuo*. The apparatus is depicted in fig. 4. The solid S was placed inside a glass tube A, and the connecting tube B joined on afterwards. This tube B led to a high vacuum pump, whilst the lower end of A was drawn out into a very fine capillary tube. A wider tube C was joined to the 1-mm. wide capillary attached to V, so that the final form of the apparatus was as indicated. The whole was then immersed in an electrically heated oil-bath and the apparatus exhausted. To facilitate the passage of the gas molecules from V to the pump it was necessary to cut small grooves around the lower edge of the cylindrical specimen S.

In this connection it is expedient to note that the oil-bath must be capable of withstanding a temperature at least 20° C. above that of the melting point of the solid. The reason for this is at once apparent, for the specimen receives heat principally by radiation only, for it is not until a mass of molten metal has formed in the lower part of A that the metal receives thermal energy by conduction. Meanwhile, the molten metal has been forcing

its way through the fine capillary into the container K. At this stage the melt is indistinguishable from a mass of clean mercury.

Experimental Details.

The metallic eutectics having been thus prepared in each case, it was necessary to break the tube of fig. 4 at X and, after the mass of metal contained in V had been determined, to attach this to the rest of the apparatus, which had been cleaned and dried in the usual manner adopted in such cases. The control of the gas pressure inside the volumenometer, which was essential during the last piece of glass manipulation, was always made through a calcium chloride tube, so that the contamination of the volumenometer was reduced to a minimum. The gases and vapours were now removed from the apparatus by using a "Hyvac" pump in conjunction with charcoal and liquid air; even after the apparatus had been allowed to stand all night no trace of leak was discernible when an attempt was made to remove any air by using a Toepler pump.

The apparatus was then filled with pure dry nitrogen in the manner referred to above, and in the first series of experiments in which the pressure difference was measured, the pressure in the apparatus was made equal to the existing atmosphere pressure by adjusting the mercury levels in the barrel of the pump and the mercury reservoir which was attached to it. In the latter type of apparatus such an adjustment was not necessary because the pressure does not occur explicitly in the formula (3).

The temperature of the bath was controlled by means of a thermostat. This consisted of a large copper vessel, heated electrically. The current through the coils was adjusted by means of a rheostat so that the temperature was a little lower than that at which it was desired to make a determination of the volume of the solid. The elevation of the temperature to a final value was made by means of a suitably controlled current through a resistance mat. This mat was fixed to an iron support, the electrical insulation being obtained by the use of mica strips. The mat was in direct contact with the oil, so that the thermal capacity of this subsidiary heating element was minimised: such a condition is essential to the maintenance of a real steady temperature.

The period in which thermal energy should be generated in this mat was controlled by means of a thermo-regulator, similar to one described by Gouy and used, at a later date, by Callendar and Barnes in their researches on the specific heat of water. Fig. 5 is a diagrammatic representation of the regulator and its adjuncts. A bulb A contains glycerol over mercury and is connected to a 1-mm. diameter capillary tube in which the mercury is displaced by the expan-

sion or contraction of the glycerol. The final level is adjusted by means of the screw S′ and a reservoir containing glycerol. Former workers, using these

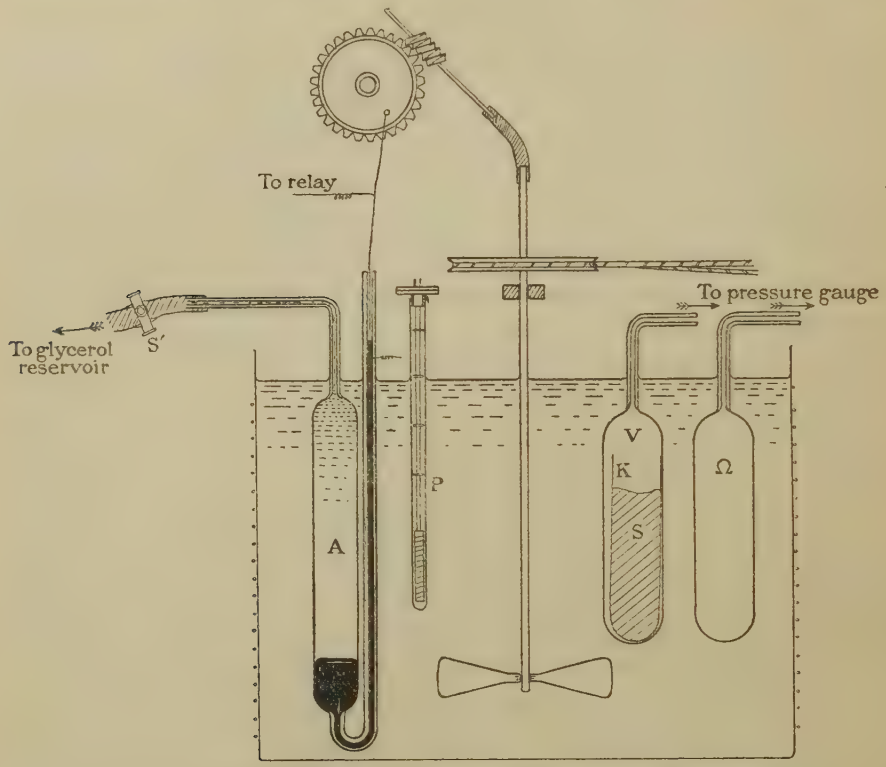


Fig. 5.—The small heating coil referred to in the text is not shown above.

thermostats at lower temperatures, have controlled this level by means of a three-way tap fixed to the capillary tube ; no such device was permissible here on account of the difficulty of discovering a lubricant which would withstand the higher temperatures to which it would have been subjected in these experiments. A platinum wire point has conferred upon it a to-and-fro motion along the capillary, the motion being effected by means of a pivot on a wheel, which was driven by means of a worm wheel operated from the main stirring shaft. The distinctive feature of a regulator of this type is that the electric control is more precise, since the mercury never sticks to the platinum point. When the relay is not excited, the current is passing in the resistance mat ; but when the relay is operated, then this coil is shorted.

At the higher temperatures the bulb A contained air only, but the control of temperature was not so good, a defect probably associated with the fact that,

when contact is made between the mercury and platinum, the pressure inside the bulb A is raised.

In such an apparatus as the one under discussion, the glass taps are a source of much trouble unless due precautions are taken. The pressure in the apparatus may be nearly two atmospheres, and thus there is a tendency for the glass stopcocks to move outwards; the effect of such a motion is manifested in the many "streaks" which appear in the lubricant on the taps. Such a defect was remedied by the use of strong spiral springs bent around the arms of the stopcock.

The volumes of the bulbs V and Ω were found by calibration with known masses of water at a definite temperature, whilst mercury pellets indicated the volumes of such portions of the capillaries as were involved in the calculation. The bulbs V and Ω had a capacity of about 60 c.c. and about 20 c.c. of metal were used. The capillaries were about 1 mm. in diameter, and in future work these will be reduced, so that the corrections will tend to become less.

The capillary tubing which was used for the gauge had a diameter of 2 mm. and was uniform in cross section. The uniformity was such that there was no difference in level when oil was introduced into both limbs, whatever the quantity of oil used, provided, of course, that the two limbs were open to atmospheric pressure.

One other source of trouble which was only apparent in the first type of dilatometer was due to the fact that the grease of the stop-cock F (see fig. 2) was dissolved by the paraffin oil, so that the oil, under pressure, tended to exude. This tap F had been included so that the oil could be placed inside the gauge after the apparatus had been filled with nitrogen. In the second apparatus there was no such tap, the oil being placed in the gauge before evacuation. The two taps L and M (see fig. 3) served to isolate the oil in the gauge during exhaustion from the remainder of the apparatus.

The density of the specimens, at room temperature, was found from observation on the apparent loss in weight when a sample was suspended in water. From the known mass of metal in the apparatus its volume was estimated.

Mode of Procedure.

When the dilatometer had been duly mounted on a stand, the thermostat was maintained at room temperature until the temperature was constant and uniform throughout, when the taps P and Q were closed, and remained so until the close of the series of experiments. The evaluation of the left-hand sides of the equation (3) was then made, and the experiment repeated at a higher

temperature, to discover the dependence of the volume of S upon temperature. Since, if accurate results are to be obtained, the temperature gradient throughout the bath, bulbs and specimen must never exceed 0·02° C., observations at any temperature were only made at intervals of never less than 90 minutes. The establishment of such steady conditions was proved by the following facts :—

(i) The resistance of the platinum thermometer was constant.
(ii) The levels of the gauge liquid did not alter.
(iii) The height of the mercury above the glycerol or air in the thermo-regulator was constant.

Typical Result for Bi-Sn Eutectic.

Initial conditions gave for the L.H.S. of equation the numerical value 0·3028. $t = 132{\cdot}3°$ C. volume of mercury abstracted $= 0{\cdot}596$ c.c. at 18° C.

$V = 41{\cdot}66$ c.c. $\quad \Omega = 47{\cdot}56$ c.c.
$v = 2{\cdot}565$ $\quad \omega = 2{\cdot}420$ at 27° C. (effect. temp.);

therefore, if S is the volume of the solid at this temperature, we have

$$0{\cdot}3028 = \frac{41{\cdot}66 - S + \frac{405{\cdot}3}{300}(2{\cdot}565) - 6{\cdot}056}{47{\cdot}56 + \frac{405{\cdot}3}{300}(2{\cdot}420) + \frac{405{\cdot}3}{291}(0{\cdot}596)},$$

whence $S = 23{\cdot}43$ c.c.

Summary of Results.

Table II.—(i) Lead-Tin Eutectic.

Temperature.	Volume.	Specific volume.
° C.	c.c.	c.c. per gm.
Series (a)—		
12·4	6·91	0·1192
153·9	7·05	0·1216
167·3	7·06	0·1218
176·0	7·07	0·1220
182·4	7·08	0·1221
Series (b)—		
15·0	17·79	0·1193
179·3	18·22	0·1222
197·2	18·68	0·1253
205·5	18·73	0·1256
186·6	18·66	0·1251
182·6	18·66	0·1251
180·0	18·23	0·1222

The results in the above table are recorded in the order in which the observations were made. Since this particular alloy does not expand on solidifying, it was possible to obtain a value for the specific volume at a temperature below the melting point (see Table II, last row). The agreement with the previously determined value in this neighbourhood is excellent, a fact which indicates the reliability of the present method.

Table III.—(ii) Tin-Bismuth Eutectic.

Temperature.	Volume.	Specific volume.
° C.	c.c.	c.c. per gm.
13·0	23·14	0·1165
58·0	23·23	0·1170
86·0	23·31	0·1174
117·0	23·40	0·1178
132·3	23·43	0·1180
135·5	23·43	0·1180
165·2	23·48	0·1182
184·1	23·55	0·1186
213·1	23·68	0·1193
151·9	23·36	0·1176
168·1	23·41	0·1179
152·6	23·35	0·1176
136·7	23·24	0·1170

Again the results are recorded in the sequence in which the experiments were conducted.

During the time when the experiments on the Bi-Sn eutectic were in progress, some observations were made which are worthy of record. After the attainment of a steady temperature throughout the bath, the pressure difference did not remain zero but became finite, the variations not being continuous but spasmodic. It was supposed that these happenings could be attributed to the development of microscopic cracks in the metal, and that when these occurred the nitrogen filled them at once so that the gauge level altered. These were avoided in the final work by annealing the metal for several hours in the apparatus at a temperature about 1° below its melting point. During this treatment the cracks were developed to such an extent that afterwards no more were formed. It can, therefore, be concluded with certainty that this method does measure the volume of metal actually present.

The results which have been obtained for these two eutectics are recorded graphically in figs. 6 and 7 ; in them the specific volume, *i.e.*, the volume of 1 gm. mass, is indicated as a function of the temperature. It is at once apparent from such curves that the function is linear for the liquid or solid metal and

that at the melting point there is an abrupt change in volume, such change taking place wholly at the melting point, and not over a range of temperature.

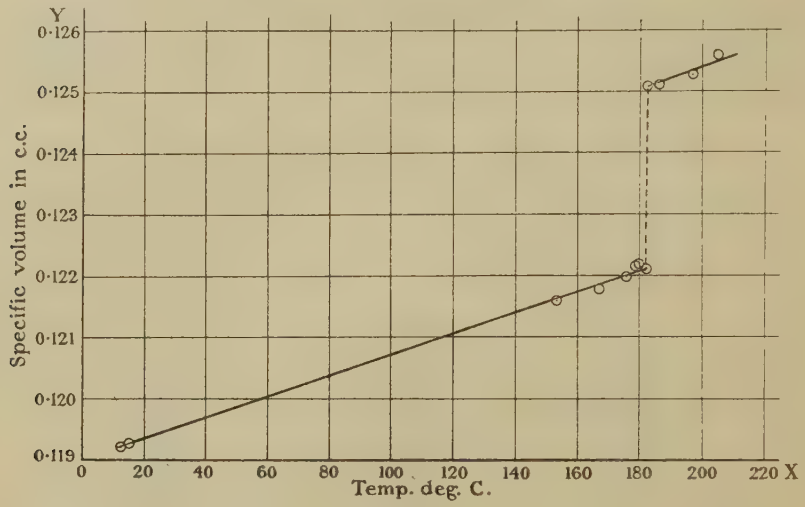


Fig. 6.

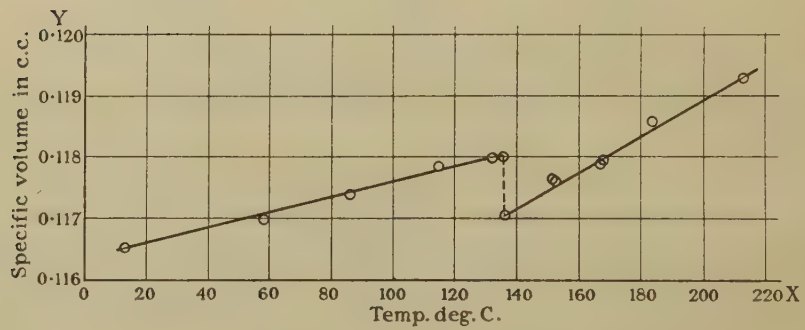


Fig. 7.

In the case of the Pb-Sn alloy the coefficient of expansion of the liquid is practically identical with that of the solid, but this is not so in the case of the Sn-Bi eutectic, where a marked difference is found. From these curves we obtain the following results :—

(i) Contraction on solidifying for Pb-Sn eutectic = 2·5 per cent.

(ii) Expansion on solidifying for Bi-Sn eutectic = 0·8 per cent.

The present method appears to be limited in its application only by the choice of suitable baths and the materials out of which the glass is made or replaced.

A criticism of the present method which is almost certain to be made is that

gas will be evolved from the glass walls and evolved or absorbed by the metals. That the volume of gas which is exuded when glass is heated is large, is only so when the gas is measured under low pressures and the temperature elevation has been considerable. Such evolution cannot be anything but infinitesimal in the present experiments. In fact, if it were great, the validity of the temperature deduced from a gas thermometer would be destroyed.

Regarding the absorption of nitrogen by the metals which have been used, it has been shown* experimentally that the absorption is exceedingly small.

Experiments with Tin.

The above apparatus has been used by Mr. W. E. Goodrich, of this department, to study the volume changes associated with the solidification of tin. The results of Table IV are deduced from the observations which he has kindly placed at my disposal. Fig. 8 records these results graphically. They are of

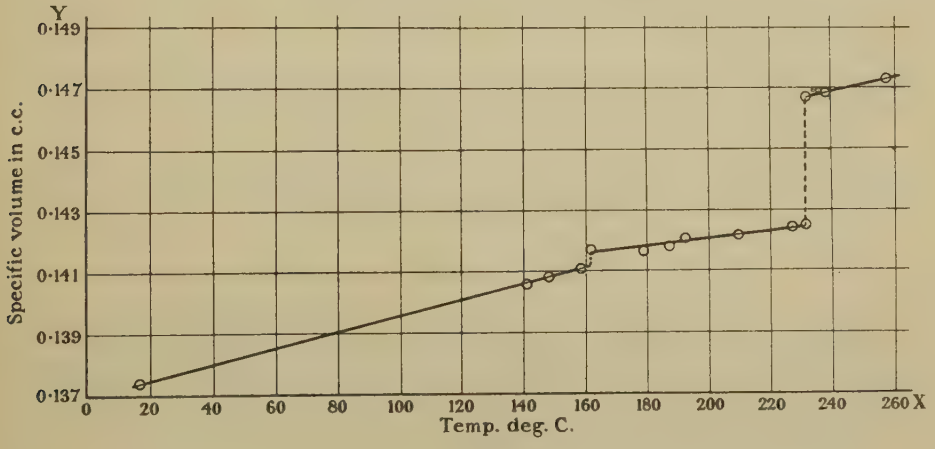


Fig. 8.

exceptional interest since they testify further the reliable nature of this new differential dilatometer. The figure proves that not only does the method measure the abrupt volume changes at the melting point, but also indicates other discontinuities which may occur. In the present case the sudden volume change at 162° C. is clearly marked.

* Iwasé, 'Sci. Rep. Tôhoku Univ., vol. 15, p. 531 (1926).

Table IV.—Tin.

Temperature.	Specific volume.	Temperature.	Specific volume.
° C.	c.c. per gm.	° C.	c.c. per gm.
16·0	0·1375	192·2	0·1421
140·4	0·1406	210·5	0·1421
148·5	0·1408	227·5	0·1424
158·6	0·1411	231·9	0·1425
162·0	0·1418	232·0	0·1467
179·0	0·1417	238·7	0·1468
187·9	0·1418	257·4	0·1472

In conclusion the author would like to acknowledge the helpful criticisms and suggestions which he has received from Prof. C. H. Desch during the course of the work.

The Cause of the Colours shown during the Oxidation of Metallic Copper.

By F. Hurn Constable, M.A., Ph.D. (London, Cambridge), Fellow of St. John's College, Cambridge.

(Communicated by T. M. Lowry, F.R.S.—Received April 26, 1927.)

During the initial stages of the oxidation of a clean metallic copper surface, and before the normal black colour of cupric oxide is finally attained, bright colours appear which pass through the first and second orders of the series of colours observed and tabulated by Newton* as characteristic of thin films of air of increasing thickness. These facts have been known for a considerable time, but though the interference colours of thin films of air and those of the oxide film are produced in the same order, grave doubts have existed as to whether interference was the cause of the colours shown by the copper.†

This problem has been approached from measurements of the intensity of the light reflected from copper oxide films of known thickness, and the results show that interference is actually the cause of the production of the colour,

* Rollet has very carefully re-investigated the phenomenon, and tabulated the colours and the corresponding thicknesses ('Sitzungsb. Akad. Wien,' vol. 67, p. 229).

† Mallock, Raman, and Hinshelwood thought that interference colours were impossible in a film backed by a good reflector.

though the result is complicated by the opacity of the film, the dispersion of the oxide, and the scattering of the light complementary to the film colour when the metal is burnished.

Owing to the high refractive index of the oxide the light within the film is almost normal to the surface inside the film whatever may be the angle of incidence in air, and thus the absence of any marked variation in the colour, with the direction of the incident light would be expected, and is characteristic of the colours produced.

The diffraction theory supposes that the colour is due to many small spheres of copper sheathed around with oxide, the extent to which each is converted into oxide determining the colour of the diffracted light.* It will be shown later that Hinshelwood's own observations strongly support the interference theory and are confirmed by the present work. The remarkable optical properties of, and the changes in colour of, thin gold and silver films observed by Beilby,† and those of potassium and sodium films on glass observed by Wood,‡ drew the attention of Maxwell Garnett§ to the behaviour of minute metallic spheres whose diameter was small compared with the wave-length of light, and he considered quantitatively the problem presented by light traversing a medium containing many small metal spheres to a wave-length of light. He found, using the results of Herz,|| Rayleigh,¶ Lorentz,** and Larmor,†† that such a medium was optically equivalent to a substance having a calculable refractive index n' and absorption κ'; and gave a satisfactory explanation of the behaviour of these thin metallic films. The results in no way resemble or give a clue to the behaviour of the copper films on oxidation.

The objections raised by Mallock and Raman* were founded on the apparent impossibility of obtaining interference in a thin film backed by a perfect reflector, for the stream of light that would be transmitted in the absence of the reflector, and which would be complementary to the colour of the film, is superposed on

* Hinshelwood, 'Roy. Soc. Proc.,' A, vol. 102, p. 318 (1923); Raman, 'Nature,' vol. 109, p. 105 (1922), and Mallock, 'Roy. Soc. Proc.,' A, vol. 94, p. 561 (1918), having raised objections to the interference mechanism of production of colours on tempered steel some time previously.

† 'Roy. Soc. Proc.,' A, vol. 72, p. 226 (1903).

‡ 'Phil. Mag.,' vol. 3, p. 396 (1902).

§ 'Phil. Trans.,' A, vol. 203, p. 385 (1904).

|| 'Ausbreitung der electrischen Kraft, Leipzig,' p. 150 (1892).

¶ 'Phil. Mag.,' vol. 44, pp. 28–52 (1897).

** 'Wied. Ann.,' vol. 9, p. 641 (1879).

†† 'Phil. Trans.,' A, vol. 189, p. 283 (1897).

the ordinarily reflected light and no interference effects can be produced. Wood* has shown that if a collodion film be deposited on a bright silver mirror by means of chemically pure ether, then no colours are developed, thus experimentally confirming the theory. It was found that if the surface was breathed upon, bright colours immediately made their appearance ; and microscopic examination showed that the plane surface of the collodion had developed granulations that were just visible under the highest powers of the microscope. Light of colour complementary to that of the film was found to be scattered sideways by the film. The explanation advanced by Wood is as follows :—

" If we consider some value of λ, for which the path difference between the rays reflected from the collodion and metal surfaces amounts to an odd number of half waves, the colour corresponding to this wave-length will be weakened in the reflected beam owing to interference. In the case of transparent thin films the absent colour appears in excess in the transmitted light, while in the present case it is thrown back through the film by the metal surface. It is thus clear that the colours which are weakened in the reflected light are made to traverse the frilled film a greater number of times than the colours for which the path difference is an even number of half waves. This accounts for the fact that these colours are more strongly scattered by the granulations of the films."

It was also found by direct experiment that frilling the collodion surface increased the brilliancy of the reflected light about three times, and that the light was reflected regularly. Such results with collodion at once recall the behaviour of copper reduced from oxide during activation. At the first oxidation the colours are dull, but become much brighter after two or three alternate oxidations and reductions. The increasingly fine structure present after activation enables the light, which owing to interference cannot be transmitted through the film, to be scattered sideways from it. It is worthy of note that a brightly polished rod of copper shows brilliant colours on its first oxidation, and that the duller the polish the duller the colour.

Microscopic† examination of the partially oxidised copper surface with a power of 1,250 showed that the copper had a granular structure. The surface was composed of patches ; each patch having the same colour had similar grains. Those areas on a composite film with the larger grains were oxidised to a much less extent than those which had a smaller size of grain. Each individual grain appeared of a uniform colour, the intensity of which was greatest at the centre

* ' Phil. Mag.,' vol. 7, p. 385 (1904).

† I am much indebted to Mr. Rawlings for the loan of his metallurgical microscope.

and faded off toward the grain edges, which seemed black. Thus the colour of the diffusely reflected light is not compound in the sense that the centre of the grain is of different colour to that of the edges. Thus the colours shown by a film may be made truly homogeneous if special precautions be taken in the preparation and activation of the films so that the oxidation of all parts of the surface occur equally fast. A method has been developed which produces such uniformity quite often.

The Production of Uniformly Activated Films.

A clean piece of linen, saturated with very viscous oleic acid, was dipped in powdered cupric oxide, and a thin coating given to the china clay rod, which appeared of a uniform grey colour. The rod was made uniformly red hot, and alternately oxidised and reduced at this temperature by taking it in and out of the gas flame. When cooling, the surface was rubbed with another oxide coated rod, until the oxide layer was polished to a uniform jet black. A second coat of oxide was put on similarly and polished, but the final layer was left unpolished, for the colours appeared more markedly on the dull surface. The hot rod, held in an ordinary cork, was inserted in a test-tube containing boiling methyl alcohol vapour. Immediate reduction to copper occurred, and the rod was activated by repeatedly inserting and withdrawing it from the test-tube. When the rod was so cool that reduction became slow, the cork holding the rod was firmly inserted in the tube, and the whole was left to cool in the methyl alcohol vapour. When cold the rod was removed and placed in a test-tube evacuated by a pump, and left for 30 minutes so that the alcohol remaining in the rod could be vaporised off. Such a film possesses a metallic lustre, and appears continuous to the eye.

If care be taken to heat the rod uniformly before the reduction then very homogeneous copper films are produced even if the treatment is not continued ; but three further oxidations and reductions were made in the furnace, fig. 1A, under standard conditions of temperature.*

The Production and Fixation of the Colours.

(1) With activated supported copper films. The china clay rod E supporting the metallic copper film, activated as in the previous paragraph, was introduced into the electric furnace, fig. 1A, being held concentrically in the furnace

* Palmer, ' Roy. Soc. Proc.,' A, vol. 103, p. 444 (1923), and Dunn, *ibid.*, vol. 111, p. 213 (1926), both agree that after two or three oxidations and reductions under standard conditions the film settles down into a reproducible state.

tube by the clip P on the non-conducting rod O, which passes through a rubber bung at the end of the tube. To secure uniformity of temperature over the whole surface of the metallic film and at the same time to render the film always visible to the operator, the glass furnace tube was uniformly wound with nichrome wire. Through the spacing between the individual turns of wire the metal film could be watched, though it is advisable not to use carbon monoxide very frequently as reducing agent after oxidation, because a slight film of volatilised copper settles on the glass making the glass turn a dirty brown colour. After the introduction of the metallic film, the glass furnace was evacuated, filled with carbon monoxide to prevent oxidation of the copper, and heated to the desired temperature.* When temperature equilibrium had been attained, the carbon monoxide was removed by evacuation, and air suddenly admitted. The colours appeared on the copper with a rapidity that varied with the temperature. It was found that the oxidation could be immediately brought to a standstill by evacuating the air from the furnace tube. The absence of leak, and the stability of the films to heat treatment were both demonstrated by heating the film in the evacuated furnace to 350° C. for one hour when not even the slightest change of colour was noticed. The current was switched off while the furnace was still evacuated. When cold, it was found that the oxide film was sufficiently stable to keep its colour in dry air for some days. Wetting the film hurried the oxidation considerably, but after a time all films so prepared reached the golden yellow stage of oxidation even at room temperature.

Stable coloured films can thus be produced for spectrophotometric analysis, by sudden evacuation of the furnace tube when the desired colour is attained, and subsequently allowing the whole to cool. It was desired to study the light from a film of gradually increasing thickness, and so the oxidation was arrested at frequent intervals, and the light from the rod compared with the light reflected under similar conditions from a white china clay rod of similar dimensions. At first attempts were made to use a single rod for one whole series of experiments, placing it back in the furnace after examination, evacuating, heating to the desired temperature, and re-admitting air for the same time interval as before. Exhaustion caused the oxidation to stop, and on cooling the colour was examined as before. It was found, however, that this pro-

* The furnace temperature was calibrated against the ammeter reading A, by inserting a thermometer in place of the rod E. Any desired temperature could be obtained by means of the rheostat R. The time taken for the tube to reach temperature equilibrium was estimated by similar means.

cedure did not produce the full colour sequence, and that though the oxidation did proceed after having been arrested, yet it proceeded much less regularly. In the end it was found far more satisfactory to use a different rod for each experiment. Care was taken to prepare each in exactly the same way, and to heat them to the same temperature for the same time before oxidation. The only difference in the treatment of each lay in the increasing time intervals they were exposed to the air. Evacuation followed immediately after, and the furnace was allowed to cool to room temperature always before the rod was removed into the air again.

(2) Massive copper was obtained in rods, of the same dimensions as the china clay, and well polished with emery paper. A single rod was introduced into the furnace in a carbon monoxide atmosphere, and when the tube had attained the desired temperature it was evacuated and air let in. The colour of the rod was fixed by evacuation, followed by rapid cooling, as before.

Arrangement of the Spectrophotometer.

It was first attempted to determine the spectrophotometric record of the film oxidised *in situ* in a glass tube surrounded by nichrome heating wires; but the attempt resulted in failure because the reflection from the glass surface was very strong and reduction of the film after oxidation caused the formation of a thin film of metallic copper on the glass, which became increasingly opaque. The method described in the previous paragraph was highly satisfactory, producing a film which could be handled easily in air. Fig. 1c shows the apparatus for comparing the intensity of the light (of any given wave-length) reflected from the film, with that reflected from a rod of pure china clay of the same dimensions. The coloured rod E was held in a screw clip, side by side with the china clay rod F, in the stand G. Light from the 400 c.p. Osram lamp A, whose filaments were 27 cms. away from both rods, shone on the two rods fully, while the fixed shield B and the moving shield C, caused the black cloth D, 1 metre in the rear of the rods, to be in deep shadow. The use of a bright source of light enabled the experiments to be carried out in the ordinary laboratory, no specially darkened room being necessary. Fig 1B shows the rods E and F backed by the black cloth, with the positions of the windows of the Nutting photometer H dotted in front of them.

Care is necessary in the adjustment of the rods E and F so that their images may be superposed by the Nutting photometer H. The distances* of E and F from the windows K and L were arranged so that the images of both would

* EK = 19·0; FL = 17·2 cms. in the actual apparatus.

coincide in the photometer H when the rods examined were vertical, and were adjusted so that the image of each in the respective window of the photometer

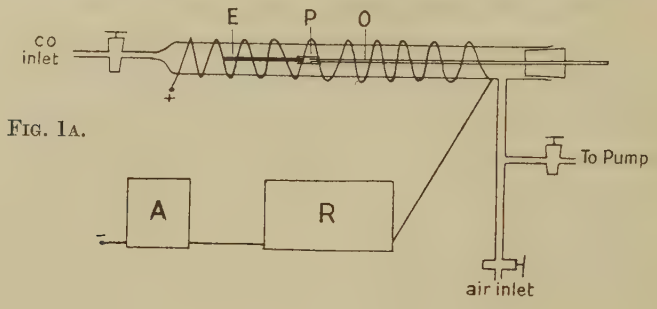


Fig. 1A.

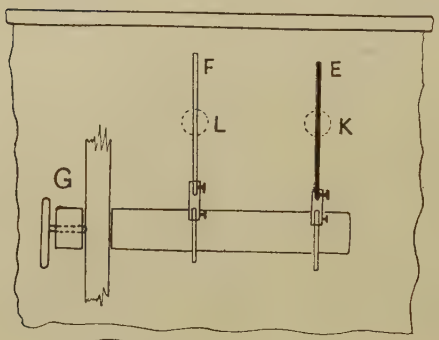


Fig. 1B.

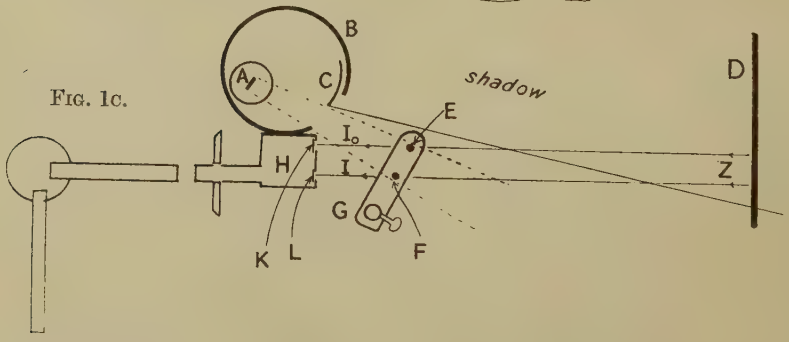


Fig. 1C.

Fig. 1.—Apparatus.

was completely hidden by the object when viewed from Z. The photometer was used in conjunction with a Hilger wave-length spectrometer, the centre of the triply divided field of which showed the spectrum of the light from F the china clay standard, while the upper and lower spectra were those of the coloured film under examination. Any portion of the spectrum could be selected for examination, and by rotation of the graduated second nicol of

the photometer H the intensity of the light from the copper film E, whose wave-length was read off the drum of the spectrometer, could be compared with that of the light similarly reflected from the china clay standard F.

A blank experiment showed that the light which entered from the cloth background was too small to be detected by the instrument when the standard of comparison was the illuminated china clay rod.

Corrections for the Zero Error of the Instrument, the difference of distance between the two rods and the photometer windows, and differences in the intensity of illumination of the two surfaces.

Let the suffix 1 denote quantities for the coloured film E, while 2 denotes the some quantities for the rod F.

Let

a_1, a_2 be the effective distances of the two rods from the photometer.

L_1, L_2 be the intensity of the light incident on both rods.

ϕ_1, ϕ_2 be the fraction of the incident light of wave-length λ which is reflected from the surfaces.

When the second nicol has been rotated so that the light from the china clay rod exactly balances that from the film for a given wave-length, the reading on the photometer drum d_λ is the ratio of the logarithm of the intensity of the light from the china clay rod I_0, divided by the intensity of the light from the film I.

Hence

$$d_\lambda = \log_{10} I_0/I, \tag{1}$$

But

$$I_0 = \frac{\phi_2 L_2}{{a_2}^2},$$

and

$$I = \frac{\phi_1 L_1}{{a_1}^2},\ *$$

therefore

$$d_\lambda = \log_{10} \frac{\phi_2 L_2/{a_2}^2}{\phi_1 L_1/{a_1}^2}. \tag{2}$$

If two similar china clay rods† be compared by the instrument in the same

* An inverse square law is assumed because the source of reflected light was quite small, compared with the distance from the photometer windows. Any power law would give (4) unaltered.

† The two rods were chosen as optically similar by trial. When reversal of the positions of E and F caused no change in the photometer readings, the two china clay rods E and F were taken as optically similar.

coloured light, then the value of d_λ' observed is found not to be zero, but can be made small throughout any series of experiments by regulating a_1, a_2 and L_1, L_2. This adjustment was always made. Let d_λ' be the value observed for the density with light of wave-length λ, comparing two optically similar reflecting china clay surfaces. Since $\phi_1 = \phi_2$ we have

$$d_\lambda' = \log_{10} \frac{L_2/{a_2}^2}{L_1/{a_1}^2}. \tag{3}$$

Subtracting (3) from (2) we obtain

$$\log_{10} \frac{\phi_2}{\phi_1} = (d_\lambda - d_\lambda'). \tag{4}$$

The spectrophotometer thus determines directly the logarithm of the reciprocal of the relative intensity of light of a given wave-length reflected from the copper oxide film, provided the corrections for the above sources of error are subtracted from the photometer readings.

Table I.—Corrections.

Wave-length.	Density on photometer.*		Mean correction d_λ'
	Rods arranged E F.	Rods reversed F E.	
4500	0·14	0·24	0·19
4600	0·13	0·23	0·18
4700	0·15	0·25	0·20
4800	0·08	0·16	0·12
4900	0·10	0·18	0·14
5000	0·07	0·13	0·10
5100	0·06	0·12	0·09
5200	0·07	0·13	0·10
5300	0·06	0·16	0·11
5400	0·08	0·14	0·11
5500	0·06	0·12	0·09
5600	0·04	0·12	0·08
5700	0·04	0·10	0·07
5800	0·04	0·10	0·07
5900	0·05	0·11	0·08
6000	0·06	0·12	0·09
6100	0·07	0·09	0·08
6200	0·06	0·10	0·08
6300	0·07	0·11	0·09
6400	0·06	0·06	0·06
6500	0·06	0·06	0·06

* The mean of three readings of the photometer was always taken, ample time being allowed for the eye to rest after each determination.

Table II.—Rollet's Table ('Sitzungsb. d. K. Akad. in Wien,' vol. 57, p. 229 (1878)).

Order of colour.	Colour of air film in transmitted light.	Thickness of air film in cms.
I	Brownish white	$1{\cdot}0 \times 10^{-5}$
	Clear brown	1·07
	Dark brown	1·16
	Red brown	1·24
	Dark purple	1·29
	Dark violet	1·35
	Dark blue	1·40
	Clearer blue to greenish	1·64
	Still clearer blue	2·35
	Pale blue green	2·45
II	Pale green	2·57
	Clear yellow green	2·72
	Clear yellow	2·82
	Golden yellow	3·00
	Orange	3·52
	Red	3·72
	Deep purple	3·87
	Violet	4·09
	Blue	4·35
	Clearer blue	4·65
	Bluish green	4·90

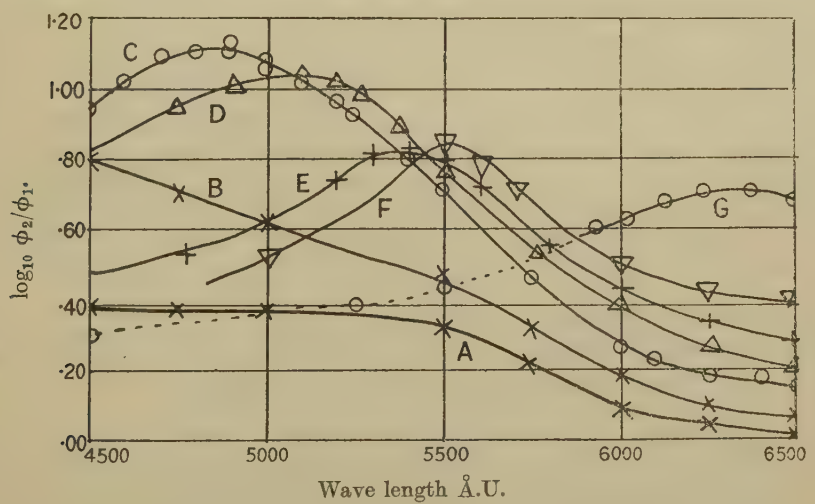


Fig. 2.—Showing absorption band moving from the violet to the red end of the spectrum as the oxide film thickens.

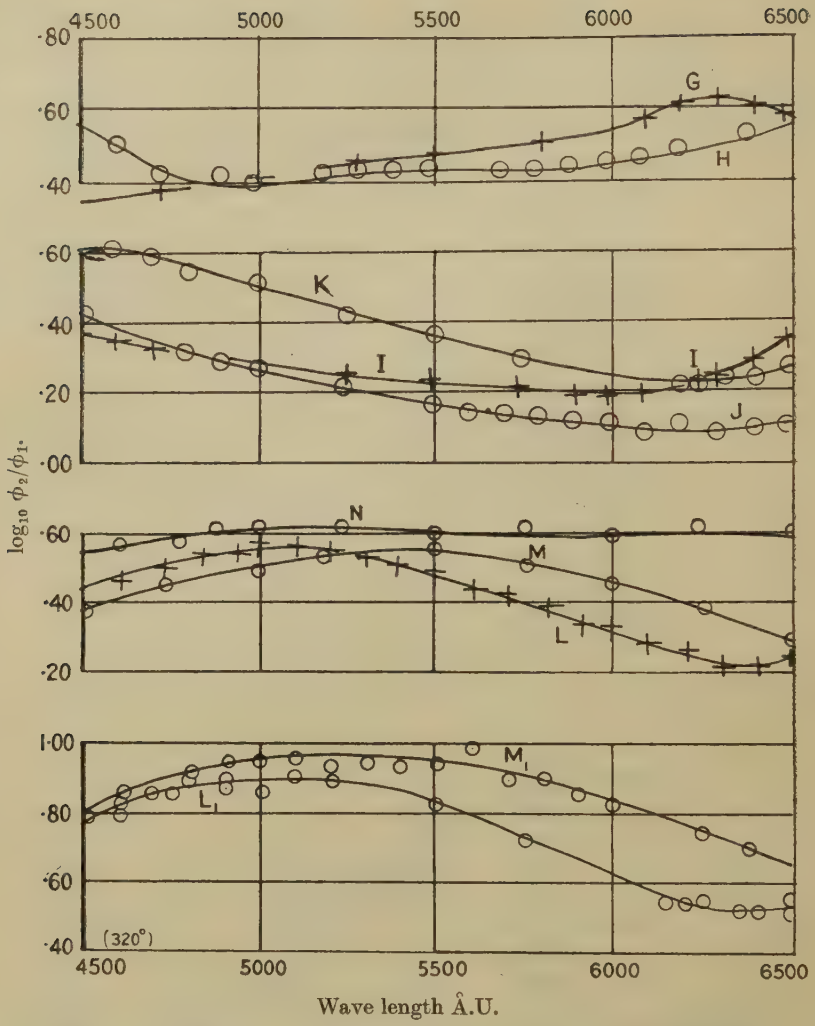


FIG. 3.—Showing the appearance of strong reflection in the violet, moving towards the red as the film thickens further ; when the band approaches the red, a second violet absorption band makes its appearance moving towards the red.

Results.

A supported copper film was oxidised for the time intervals stated below, and then allowed to cool *in vacuo*. The spectrophotometric record of the film compared with china clay was then taken, the colour of the film being recorded in accord with Rollet's table of colours given above.

Table III.

Time of oxidations in secs. at 250° C.	Colour of film (by reflected light).	Thickness of the air film of the same colour (by transmitted light).	Intensity curve of reflected light.	Wave-length of maximum	
				Absorption.	Reflection.
0	Metallic	—	Fig. 2 A	—	—
6	Brownish	1·07—1·16	C	$4·8\times10^{-5}$	—
7	Red brown	1·24	D	5·1	—
8	Dark purple	1·29	F	5·4	—
9	Dark violet	1·35	F	5·5	—
12	Dark blue	1·40	G	6·4	—
During a considerable period of time the colour is greenish but reflects light with increasing intensity as the time proceeds.					
45	Bluish green	2·45—2·72	Fig. 3 H	—	4·9
61	Yellow	2·82—3·00	I	—	6·1
75	Orange yellow	3·00—3·52	J	—	6·2
105	Orange red	3·52—3·72	K	4·6	6·3
112	Reddish	—	L	5·1	6·4
125	Feeble violet	—	M	5·5	—
150	Almost black	—	N	—	—

These results show quite clearly that the first result of oxidation is to diminish the quantity of light generally reflected from the surface. The diminution is most marked in the violet; then as the film of oxide thickens, the wave-length which is least reflected moves toward the red end of the spectrum. The greens persist over a considerable range of thickening of the film, during which the intensity of the reflected light of all wave-lengths increases, a violet wave is most strongly reflected at first, and as the film thickens the colour of the rays which are most strongly reflected move toward the red. As the red is reached an absorption band makes its appearance in the violet again and travels toward the red. Thus the film shows all the characteristics of interference colours.

Evidence that the equivalent air thickness of the Oxide Film is directly proportional to the Quantity of Metal Oxidised.

A quantity of copper wire, which had been cleaned with emery paper, was inserted into a glass tube, provided with a water manometer, which communicated with a similar glass tube into which no copper was placed. Both tubes could be evacuated or filled with air simultaneously. They were heated up to 250° C., and air suddenly admitted to both by means of a single three-way tap.*

* A blank experiment admitting pure nitrogen showed no change in the water levels on standing.

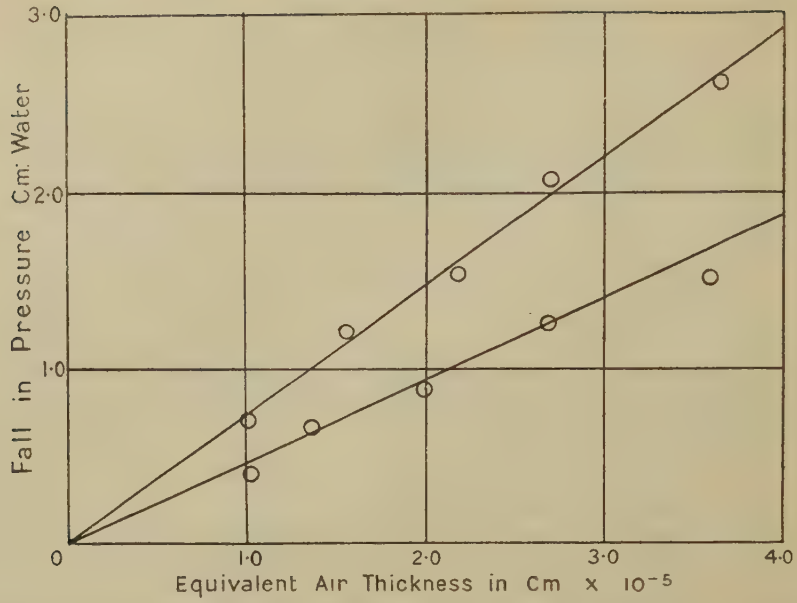


FIG. 4.—Showing the equivalent air thickness of the oxide film directly proportional to the oxygen present.

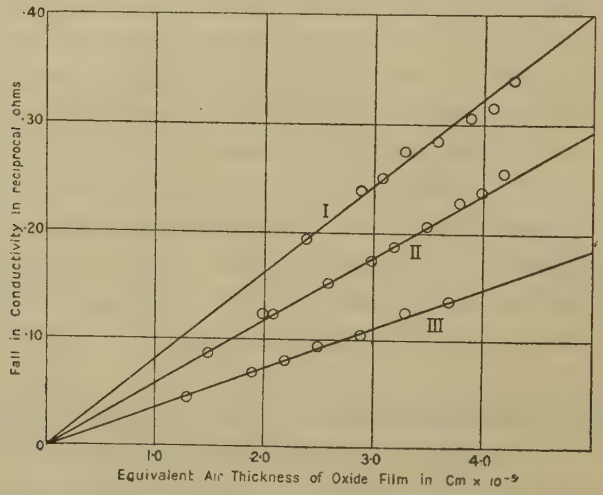


FIG. 5.—Showing the fall in electrical conductivity of an oxidised film is directly proportional to the equivalent air thickness.

The tap was immediately afterwards turned so that the two tubes were disconnected. The colours of the copper wire surface were observed and the corresponding pressures on the water manometer were recorded. The results, fig. 4, show that the thickness of the oxide film is approximately proportional to the fall in pressure in the tube, and thus to the quantity of oxide present.

Experiments were also made on supported films connecting the fall in electrical conductivity with the equivalent air thickness of the oxide layer. These are shown in fig. 5. Curve I represents a fully activated film, while curves II and III were obtained after sintering the original film by heat treatment. In all cases the thickness of the film is directly proportional to the fall in conductivity.

Theory of Production of the Colours.

The yellow metallic colour of the copper becomes brownish when the thinnest films of oxide are formed upon it, and the maximum absorption commences in the violet. Thus there is a change of phase of the light on reflection at both the film and at the copper surface making a total change of phase of one wave-length. The colour of the film is thus the same as if there were no phase change at all. Cupric oxide shows normal dispersion* which is considerable, and exaggerates the colours produced by interference.

Interference between the rays reflected at the air-cupric oxide interface, and those returning after reflection at the copper-oxide interface occur when

$$2t = (2n - 1)\frac{\lambda}{2\mu},$$

where $n = 1, 2, 3$, etc. (the value of n being limited by the general absorption in the film); and

t = the thickness of the oxide film, λ the wave-length of maximum absorption in the spectrum of the reflected light, and μ is the refractive index of the film for this coloured light. Thus

$t = \lambda/4\mu$ for the first absorption band, $= 3\lambda/4\mu$ for the second, $= 5\lambda/4\mu$ for the third; there being maximum interference in the film for these thicknesses.

For the waves to reinforce $2t = n\lambda/\mu$ where $n = 1, 2, \ldots$ hence $t = \lambda/2\mu$, etc., for maximum reflection.

Proof that the absolute thickness of the Oxide Film is of the order of and approximately the absolute magnitude of that required by the Interference Theory.

If a piece of burnished metallic copper be oxidised to a given colour, and measurements are made of the mass of cupric oxide† per unit area and the

* Kundt, 'Sitz. Akad. Wiss.,' Berlin, p. 268 (1888). Owing to the high refractive index of the cupric oxide the light inside the film is almost normal to the surface.

† Metallurgists are agreed that though cupric oxide is stable at low temperatures, and cuprous unstable, the bulk of a thick oxide film is cuprous oxide. The surface, however, was invariably black at the finish of these oxidation experiments, and it was assumed that the colours occur in a thin veneer of cupric oxide. A thin coating of cupric oxide over the cuprous has been observed in the majority of cases of oxidation.

position of the interference or reflection band in the spectrum is determined, then if the refractive index of the oxide be known the thickness of the oxide film can be calculated by two independent methods.

Table IV.

Order.	Colour.	Position of maximum absorption in spectrum λ.	Refractive index corresponding to absorption band μ.	Mass CuO per sq. cm. of surface in gms.	Thickness of film.	
					$\lambda/4\mu$ and $3\lambda/4\mu$.	$x/6\cdot3$.
		cms.		gms.	cms.	cms.
I	Red brown	$4\cdot8\times10^{-5}$	3·06	$3\cdot1\times10^{-5}$	$0\cdot39\times10^{-5}$	$0\cdot48\times10^{-5}$
I	Dark violet	5·5	2·89	3·3	0·48	0·52
II	Orange red	4·8	3·10	9·1	1·18	1·45
II	Blue	5·8	2·82	11·4	1·54	1·81

The values of the refractive index assumed are those obtained by plotting Kundt's data* and drawing a straight line through the points. The value of the refractive indices taken in the table above were so obtained, but as can be seen by reference to the paper they can be regarded as approximations only though sufficiently accurate for our purpose. A copper prism whose angle was 13 degrees, was ignited, and the deviation produced with light whose wavelength was 4·3 and $6\cdot46\times10^{-5}$ cms. was measured, giving 3·18 and 2·65 for the respective indices.

The agreement between the two methods seems on the whole to be satisfactory considering the great experimental difficulties in the way of making accurate measurements.

Hinshelwood* has brought forward experimental evidence in favour of a diffraction mechanism of production of the colours, and his experimental work is practically all confirmed by the present work, but these results can nearly all be predicted from the interference theory.

Hinshelwood's Observations.

He states (*loc. cit.*, p. 322) that the brilliancy of the colours increases rapidly on activation of the reduced copper film. This was confirmed, but the colour sequence, though fainter, did not appear to be altered. Copper wire, cleaned with emery paper, showed the same colour sequence as activated reduced copper, and the colours were practically as brilliant, persisting till the third order was attained.

* *Loc. cit.*

The ratio of the amounts of oxygen combined for various surface colours was found to be constant, and the ratios that Hinshelwood determined are given below and can be seen to be in satisfactory agreement with those calculated from Rollet's values of the corresponding equivalent air thicknesses divided by the respective refractive indices.

Table V.

Colour.	Ratio of oxygen absorbed. (Hinshelwood.)	Rollet's air film thickness for same colour. Mean q.	Refractive index. μ.	Ratio of oxygen abs. on interference theory. q/μ.
Purple	0·40, 0·38, 0·41	1·29	2·97	0·47
Blue	0·64, 0·59	1·87	2·87	0·71
Light green	1·00	2·72	2·97	1·00
Second purple	1·3	3·87	3·0	1·4
Second blue	1·6	4·63	2·9	1·7

Table VI.—The values of the mass of cupric oxide on copper foil and the corresponding colour afford a means of estimating the film thickness corresponding to the colours in Hinshelwood's experiments.

Colour.	Mass CuO per unit area × grams. (x.)	Equivalent air thickness. Range.	Equivalent air thickness. Mean q cms.	Thickness of film. $x/6·3$.	Thickness of film. q/μ.
Purple	$3·0\times10^{-5}$	1·29	$1·29\times10^{-5}$	$0·47\times10^{-5}$	$0·43\times10^{-5}$
Blue	4·3	1·40—2·35	1·88	0·67	0·65
Light green	7·3	2·72	2·72	1·16	0·95

The values of the thickness thus obtained are of the same order of magnitude, and are numerically quite as close as would be expected. When the film is activated by alternate oxidation and reduction the amount of oxygen taken up by unit area of the film is increased, but the effect may be explained by supposing that the activated copper has a granular structure so that its surface is increased. Thus this evidence is consistent with the interference mechanism of production of the colour. The table below shows the ratio of the plane surface area of the copper to the area calculated from the amount of oxide present on the surface and the observed colour.

Table VII.—The increase in area of copper foil after heating in air and reducing.

Colour.	Ratio of plane surface of copper to surface exposed to oxidation.	
	Foil.	Foil heated for 10 hours and reduced.
Purple	1·01	2·62
Blue	1·04	2·78
Light green	1·28	2·44
Mean	1·11	2·61

Thus it can be seen that while the observations used previously to support the diffraction theory of the colours, equally support the interference theory, the behaviour of the reflected light from the film makes certain that the colour is produced by a true interference mechanism, and the absolute thickness of the film calculated from the amount of oxide formed and the position of the absorption band together with the refractive index agree. The general absorption in the film is such that colours after the second order are difficult to obtain, and though third order colours can sometimes be observed, often the film merely becomes black.

Discussion of Results.

Objections have been raised to the interference theory of colour by pointing out that though the colours are formed in the order that interference colours are usually observed, when the film is afterwards made thinner the colours do not change in the reverse order.* There seems to be a random factor associated with the removal of the oxide layer, which sometimes enables the colour sequence to be retraced and at other times not. After the green stage is passed it is very rarely that the colours are reversed when the oxide is dissolved off with dilute hydrochloric acid, or if the film be reduced by carbon monoxide.

It is known that carbon monoxide only reduces copper oxide at the interface between the oxide and the metal, and that the velocity of such chemical change does vary with the energy excess at the interface and the permeability of the oxide film.

* U. R. Evans has obtained a considerable amount of qualitative evidence in favour of the interference theory, and has devised a method of cathodic reduction which may be used to cause a uniform thinning of the film, and so causes the reversal of the colour sequence. 'J. Chem. Soc.,' vol. 127, p. 2484. 'Roy. Soc. Proc.,' A, vol. 107, p. 228 (1925).

Thus there is no certainty that the copper nuclei present at the base of the film will grow uniformly beneath the oxide film, and so reduce the thickness of the oxide layer in the uniform manner which was so characteristic of its growth. The appearance of the film shows that the metallic copper often appears in patches over the surface. With activated films the colour becomes yellower till a metallic tint is reached. More rarely the green film becomes darker, passing through red and brown to yellow, showing the regular thinning of the film.

The colour sequence can also sometimes be retraced if the copper film, oxidised to the green stage, be placed in cold and very dilute hydrochloric acid, when it will be observed that a dark violet followed by a red colour will be produced before the oxide is completely dissolved off, and the matte surface of the underlying copper is exposed.

In both cases it seems that it is only rarely that the chemical reagent attacks the film sufficiently regularly to cause the retracing of the colours in order. The irregular removal of the oxide in patches is accompanied by a general dulling of the film colour which gradually merges into yellow.

The production of the oxide film on metals occurs by diffusion through a layer of oxide already present, and since the volume of the oxide is somewhat greater than that of the metal from which it was produced, both processes tend to the production of a uniform film. A local thin place in the film will be supplied with oxygen more readily, while a projection will be correspondingly badly supplied. Microscopic examination of the film failed to reveal any absence of colour in the single grains, the only change in colour under a power of 1,250 being caused by the darker boundaries of the grains. Thus it appears that the spectrophotometric observations of the reflected light from the film gives the actual colour of the centre of the grains, which enables the actual colour of the thin film to be followed as the thickness increases.

My thanks are due to Prof. T. M. Lowry, F.R.S., for the use of his spectrophotometer without which the work would have been impossible.

Summary.

Chiefly owing to the apparent difficulty of obtaining interference colours in thin films backed by a perfect reflector, Mallock, Raman, and Hinshelwood have abandoned the interference theory of the formation of colours on steel and copper during oxidation. Wood has shown that collodion films deposited on bright metallic silver do not show colours, unless the collodion-air interface be made irregular in such a manner that the light which has been prevented from

emerging from the film because of interference can be scattered sideways by the film. Very bright colours may be produced in this manner.

A considerable body of evidence has been obtained that the production of the colours on metallic copper is due to interference in a thin layer of oxide. The order of production of the colours corresponds with the order tabulated by Rollet for the interference colours of air films of increasing thickness seen by transmitted light. At 210° C. purple of the second order is succeeded by black, but at 320° C. colours of the third order can be obtained before blackness supervenes. For the purpose of experimental study the colours were defined by the equivalent air thickness, *i.e.*, the thickness of an air film which would produce the same colour as the copper oxide film. It was found :—

(*a*) The fall in electrical conductivity of the oxidised film supported on china clay was proportional to the equivalent air thickness of the oxide film.

(*b*) The equivalent air thickness was proportional approximately to the mass of oxide formed.

(*c*) The wave-length of the maxima in the absorption or reflection bands in the spectrum of the light reflected from the film, move towards the red as the thickness increases. The first order colours are characterised by absorption in the violet travelling to the red. The red absorption maximum is much less marked than that in the violet and the corresponding blue colour is feeble. A reflection maximum then makes its appearance in the violet and travels toward the red, the colours all being very light. A weaker violet absorption makes its appearance as the reflection in the red is disappearing, and the colours become much duller as the absorption band moves down the spectrum, blackness becoming marked before the red is actually reached.

(*d*) The absolute thickness of the cupric oxide film calculated from the density, and the mass of oxygen taken up per square centimetre of the surface, agrees to within 30 per cent., even in the most divergent case, with the thickness calculated from the position of the absorption band and the corresponding refractive index.

(*e*) The amount of light scattered from a dull film seems negligible, but if the copper be highly polished then the scattering is marked from green films.

On the Electric Moment of the Sulphur Complex.

By A. M. Taylor and E. K. Rideal.

(Communicated by Sir Joseph J. Thomson, F.R.S.—Received May 12, 1927.)

Broadly speaking, three types of crystal structure may be recognised: (*a*) interpenetrating arrangements of atomic and electric lattices, the electrons being symmetrically disposed with reference to the atoms; (*b*) neutral lattices, of which sulphur and phosphorus among the elements are examples; (*c*) ionic lattices, to which class most salts belong, in which the electrons are unsymmetrically arranged so as to be in close relationship with particular atoms. Sir J. J. Thomson pointed out that by a decrease of the closeness of this association between atoms and the linkage electrons, this type merges into the first.* Franck† similarly distinguishes between three kinds of linkage: (i) where the electrons remain in unchanged quantum orbits, *i.e.*, homopolar linkage due to van der Waal's forces owing to mutual polarisation of the atoms; (ii) where the linkage electrons occupy new quantum orbits which may be binuclear, *i.e.*, homopolar linkage such as occurs in molecules of many gases; and (iii) true heteropolar linking where electrons have been bodily transferred from one atom to the other, and occupy new quantum orbits around a single nucleus. Linkages of type (i) clearly produce crystal lattices of class (*a*), usually possessing electric conductivity, while those of class (ii) are concerned in the formation of insulators having neutral lattices of type (*b*). True homopolar linkage should give no active frequencies of vibration, or, if any exist, they should be, according to Schaefer,‡ of extremely weak resonant amplitude, and should lie in the region where the wave-length of the radiation is comparable to the atomic spacing.

Nernst§ suggests that all crystals are built up by homopolar linkages, but this would appear to be contrary to experience, crystals generally showing sharply defined absorption maxima in the infra-red. Kossel is quoted by Bragg|| as supposing the CO_3 group in carbonates to be held together by heteropolar attraction between the ionised atoms of the group. De,¶ however,

* Sir J. J. Thomson, ' The Electron in Chemistry.'

† J. Franck, ' Trans. Farad. Soc.,' vol. 21, p. 536 (1926).

‡ Schaefer, ' Forts. d. Mineralogie,' vol. 9, p. 31 (1923).

§ Nernst, ' Z. Angew. Chem.,' vol. 36, p. 453 (1923).

|| Bragg, ' Roy. Soc. Proc.,' A, vol. 105, p. 370 (1924), and vol. 106, p. 346 (1924).

¶ De, ' Trans. Chem. Soc.,' vol. 115, p. 127 (1919).

considers the binding to be homopolar, while Bragg and Chapman,* Chapman, Topping and Morral,† Chapman and Ludlam,‡ Kornfeld,§ and Lennard-Jones|| all agree in supposing the group to consist of ions having effective charges modified by deformation of the electronic shells. Fajans and Joos,¶ on the other hand, state that, contrary to Kossel's assumption, there seems to be little reason for supposing the internal binding of the CO_3 ion to be by heteropolar linkages, they themselves being inclined to think that the setting up of an induced electric moment in the oxygen atoms, by reason of the deformability, provides a sufficient explanation of the stability of the group.

It thus appears that there is a possibility of homopolar linkage being accompanied by such deformation of the atoms concerned as to give rise to a pseudo-heteropolar linkage with consequent electromagnetic vibration frequencies. There will be an infinite gradation from true homopolar linkages giving crystals with small deformability of the atoms, through types such as that of the carbonate group, in which the deformability being large, electromagnetic frequencies make their appearance, to true heteropolar linkages such as that between the CO_3 group and the metallic ion, or between the K and Cl in potassium chloride.

Since the forces of cohesion, adsorption,** and possibly of catalytic action are apparently related to the electric moments of the molecules concerned, it is of great importance to attempt an evaluation of this quantity characteristic of the fundamental group, and, further, by a comparison of the allotropic forms, if any exist, to deduce the nature of the complex in which the electric moment is resident. To this examination sulphur lends itself particularly well. Sulphur, and, in fact, all the elements with more than four valency electrons, can combine in pairs of atoms to form electrically neutral groups,†† but that these groups in rhombic sulphur possess an electric moment is indicated by the presence of absorption bands in the infra-red transmission spectrum, examined by Coblentz‡‡ and later by Martha Schubert,§§ while the large value of the energy necessary to disrupt the crystal lattice to form the vapour phase points to the same conclusion, being of the order of 30,000 cals. per gm. molecule of S_2, or

* Bragg and Chapman, ' Roy. Soc. Proc.,' A, vol. 106, p. 369 (1924).
† Chapman, Topping and Morral, ' Roy. Soc. Proc.,' A, vol. 111, p. 25 (1926).
‡ Chapman and Ludlam, ' Phil. Mag.,' vol. 50, p. 822 (1925).
§ Kornfeld, ' Z. f. Physik,' vol. 26, p. 205 (1924).
|| Lennard-Jones and Dent, ' Roy. Soc. Proc.,' A, vol. 113, p. 673 (1926).
¶ Fajans and Joos, ' Z. f. Physik,' vol. 23, p. 1 (1924).
** Lorenz and Landé, ' Z. f. Anorg. Chem.,' vol. 125, p. 47 (1922).
†† Sir J. J. Thomson, ' The Electron in Chemistry.'
‡‡ Coblentz, ' Carneg. Inst. Wash. Pub.,' vol. 35, p. 66 (1905).
§§ Martha Schubert, " Doctorate Dissertation," Breslau, 1915.

more than double that required for the vaporisation of water, water being well known to be a strongly polar substance with a large electrical moment.

It was therefore thought to be of great interest to examine in detail the various allotropic forms of sulphur and also the liquid phase for the existence of electromagnetically active vibrations, and to measure accurately the wave-lengths at which maximum absorption of infra-red radiation occurs, and also the value of the extinction coefficient at the maxima. From these experimental quantities the magnitude of the electric moment p may be determined, and a comparison of results found in the several solid and liquid phases leads to a suggested structure for the sulphur complex.

Experimental Details.

The apparatus consisted in an Adam Hilger spectrometer with a rock salt prism for use in the region 4–15 μ and a quartz prism from 1–4 μ. The radiation was provided by a Nernst glower running in series with a lamp resistance, off a 220-volt supply battery. The image of the filament was focussed on to the aperture C in fig. 1, which could be covered at will by a sliding shutter

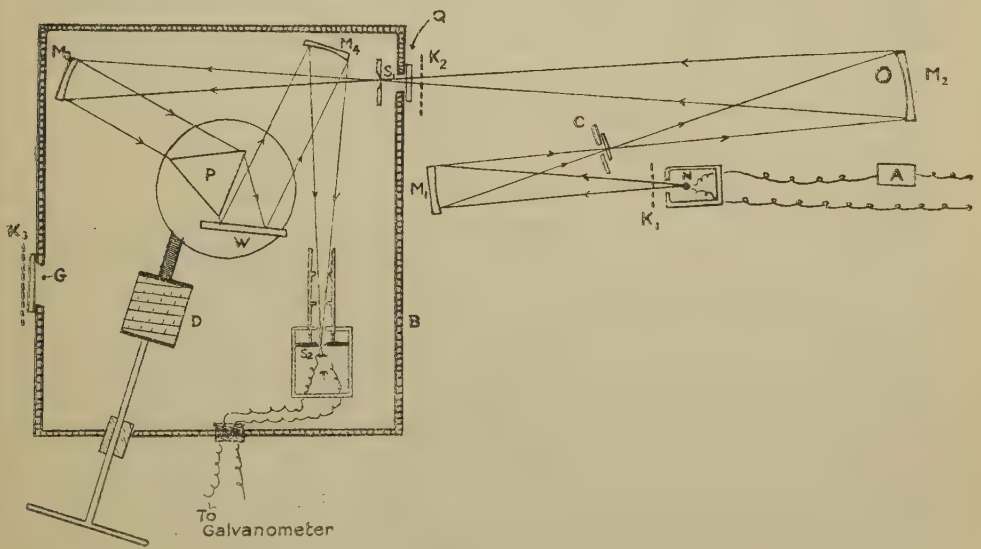


Fig. 1.—Apparatus.

N, Nernst glower.	B, Cast-iron box.
M_1 M_2 M_3 M_4, Mirrors.	Q, G, Windows.
W, Wadsworth mirror.	C, Crystal holder.
P, Prism.	K_1 K_2 K_3, Shutters.
S_1 S_2, Slits.	D, Wave drum.
T, Thermopile.	A, Ammeter.

——←—— Path of the radiation.
——→——

bearing the crystal specimen under examination, and the radiation transmitted was focussed on to the slit of the spectrometer. After analysis by the prism, the radiation was received by a sensitive thermopile connected to a galvanometer of the Paschen type having a figure of merit of about 19,000, constructed for us by Mr. A. C. Downing,* with the kind consent of Prof. A. V. Hill. The resistance was 10 ohms and a deflection of 1 mm. at 1·50 metres was produced by a current of 5×10^{-11} amperes. The spectrometer and thermopile were enclosed in a cast-iron box, which was air-tight, and which was provided with the necessary windows, etc. (one being of rock salt for the admission of the radiation). The spectrometer drum was rotated from without by a glass rod carrying a key fitting on to the axis of the drum, the glass rod passing through a rubber bung and being lubricated with a mixture of graphite and glycerine. The leads from the thermopile were sealed into glass tubes passing through a rubber bung. The reason for these precautions was that the thermopile was very sensitive to small heating and cooling effects due to chance adiabatic variations in pressure of the atmosphere,† and as the construction of the apparatus precluded the enclosing of the thermopile in a separate hermetically sealed container, the above method of shielding from atmospheric changes of pressure was adopted. It also had the advantage of keeping the spectrometer and prism at an even temperature, and a thermometer was fitted inside the case to record the mean temperature. The refractive index of rock salt varies slightly with the temperature,‡ and the prism was calibrated at 18° C. The galvanometer was stood upon a concrete pillar let into the foundations of the building, and it was shielded from magnetic disturbances by a Mu-metal shield and a surrounding soft iron case. In spite of these precautions it was found necessary to do all work after 10 p.m., and the most accurate readings were taken between 12.30 a.m. and 3 a.m., mechanical disturbances being usually at a minimum at that time. It was possible to read the deflection to 0·2 mm., and the magnitude was usually adjusted to be between 3 and 15 cm., at a scale distance of 1·5 metres. The entrance and exit slits of the spectrometer were always of equal width, and throughout the spectrum the slit widths were altered so as to use the narrowest slit consistent with accuracy of observation of the deflection.

When the slit width is expressed in terms of wave-length separation, it is clearly useless to crowd the points of observation much more than four per slit

* A. C. Downing, 'J. Sci. Inst.,' vol. 3, p. 331 (1926).
† A. M. Taylor, 'Nature,' vol. 117, p. 892 (1926).
‡ Liebreich, 'Deut. Phys. Ges.,' vol. 13, p. 1 (1911).

width, and at regions where the greatest accuracy was demanded, *e.g.*, at the bottom of an absorption maximum, this was held to be the maximum useful limit to the density of points of observation.

The Nernst glower was usually run on a current of about 0·7 ampere, and the greatest care was exercised to maintain the same value throughout a series of readings. Measurements were made by finding the deflection of the galvanometer at any particular setting of the wave-length drum (*a*) when the full radiation from the filament was entering the spectrometer, (*b*) when the radiation was partially absorbed by the specimen. The ratio of these deflections b/a was taken as a measure of the ratio I_t/I_0 where I_0 is the intensity of the incident radiation and I_t that of the transmitted radiation of wave-length lying between wave-length $\lambda-\delta\lambda$ and $\lambda+\delta\lambda$, where $2\delta\lambda$ is the slit width expressed in terms of wave-length separation.

The wave-length scale was determined from Paschen's figures in the tables of Landolt and Bornstein for the refractive index of rock salt at 18° C. In order to eliminate errors due to faults in the screw of the drum, etc., the following method was adopted for calibration. The exit slit of the spectrometer was illuminated and the parallel beam reflected from the Wadsworth mirror was received by a telescope mounted on a table, the rotation of which could be read to 5 seconds of arc. Readings were taken every 0·05 revolution of the drum, and the refractive index of the prism for radiation passing at each position, was calculated from the formula for minimum deviation $\mu = \dfrac{\sin\frac{1}{2}(A+D)}{\sin\frac{1}{2}A}$.

Paschen's figures were plotted giving a graph of the refractive index with wave-length, and a smooth curve was drawn through the points (Ruben's figures were added for comparison). From this graph and the calculated values of the refractive index a second curve was constructed showing the wave-length corresponding to each position of the spectrometer drum. This curve showed a pronounced periodic irregularity every revolution, and a smaller disturbance every half revolution, due to errors in the bearing surfaces of the axis of the drum and the lever arm of the prism table. The calibration curve so obtained agreed roughly with the maker's calibration (which was approximate). The two absorption maxima of NH_4Cl at 3·24 μ and 7·07 μ (corrected value*) found by Reinkober† and the maximum shown by calcite at 8·38 μ found by Schaefer,‡ as well as the maximum of emission of CO_2 from a bunsen burner

* Reinkober, 'Z. f. Physik.,' vol. 35, p. 179 (1926).
† Reinkober, 'Z. f. Physik,' vol. 5, p. 192 (1921).
‡ Schaefer, 'Z. f. Physik,' vol. 39, p. 648 (1926).

at 4·40 μ, were found to lie on the calibration curve, which showed proof of its accuracy.

The crystals of rhombic sulphur were cut perpendicularly to the acute bisectrix, but the direction of the prismatic sulphur was such that the normal to the plane of the crystal made an angle of 9° with one optic axis. Fairly large slabs of monoclinic sulphur (3 or 4 mm. × 6 or 8 mm.) were grown by slow cooling, in an air oven, of freshly melted sulphur* and consisted of adjacent twins,† the plane of the slabs being parallel to the basal plane "C." These slabs transmitted nearly the whole of the incident radiation of which the wave-length lay in region of zero absorption, the surfaces being sufficiently good to ensure scarcely any loss of light by scattered refraction. It was found impossible to keep these crystals sufficiently long to complete any measurements at room temperature, and it was necessary to construct a small heating coil wound on mica, which surrounded the crystal, the latter being secured to a small piece of aluminium, slotted to transmit the radiation, and thermally insulated from the holder by a sheet of mica (also slotted). By careful adjustment of the heating current, which was obtained from a constant voltage battery in series with a lamp resistance, it was possible to keep the crystal at a temperature just above the transition point of 96·4° C. for an indefinite time. Higher temperatures than this could not be used, otherwise the additional rise in temperature on passage of the radiation was sufficient to melt the crystal. For the examination of rhombic sulphur at 96° C., the current in the heating coil was slightly decreased in strength, so that the passage of radiation through the specimen should not raise the temperature so much above the transition point as to cause the crystal to break down into the monoclinic variety.

In the case of the supercooled liquid sulphur, or plastic form, it was very difficult to obtain a good preparation, or to keep it unchanged while measurements were made. For this reason the measurements of the absorption of the supercooled liquid are not exceedingly accurate. The specimens were made by dropping boiling sulphur on to a small slot 1·5 mm. × 6 mm., cut in a piece of heavy brass. The metal chilled the sulphur sufficiently rapidly for it to remain supercooled for half an hour or so, and if the brass rested upon a piece of mica during the preparation, it was occasionally possible to strip off the mica, leaving

* This particular form of prismatic sulphur appeared to grow much more readily from a melt of fresh roll sulphur than from sulphur which had already been melted and cooled a number of times, and appeared to be dependent upon the relative proportions of Sλ and Sμ in the molten liquid.

† Groth, "Chemische Kristallographie," I, p. 29.

a clean bright surface of the plastic sulphur, thus minimising the loss of energy by scattering. Other methods of preparing plastic sulphur on a support, *e.g.*, mica, rock salt, etc., were tried but abandoned, owing either to the difficulty of supercooling or to the opacity of the support.

In order to make measurements with liquid sulphur, a small cell was constructed, having 0·3 metre of fine platinum wire wound upon a small glass former of U-shape, about 3·5 mm. by 9 mm. This former was cemented between two rock-salt plates 0·5 mm. thick, with Ash's dental cement, and the cell so produced was cemented to a brass plate having a slot 8 mm. × 3 mm. cut in it ; the cell was thermally insulated from the brass by a packing of mica, and the cement was varnished with collodion to prevent any leakage of the molten sulphur. The resistance of the cell was about 4 ohms and a current of 1·06 amperes was sufficient to maintain the cell at a temperature of approximately 125° C., in which condition the cell could be filled with molten sulphur. The thickness of the layer of liquid sulphur thus available for measurement was 0·20 cm.

In all cases where the specimen was above room temperature, slight emission was observed from the sulphur, and this was eliminated by opening the shutter of the entrance slit of the spectrometer before reading the zero of the galvanometer, and then on opening the shutter which covered the source of radiation the galvanometer deflection was proportional to the energy transmitted. This emission was so small that only in one case, liquid sulphur at 125° C., was it found possible to plot a curve showing the emission bands.

Table I.—Chief Maxima in Absorption Spectrum of Sulphur.

Form.	Temperature (approx.).	Wave-length in cm. × 10^{-4}.		
	°C.			
Rhombic	18	7·76	10·73	11·90
	96	7·79	10·80	11·97
Prismatic	96	7·79	10·80	11·97
Plastic	18	7·68	—	11·83
Liquid	125	7·74	10·78	11·90

Emission Maxima.

	°C.			
Liquid	125	7·75 and 8·50	10·75	12·00

Table II.

Wave-length.	Substance.	Temperature.	Galvanometer deflections.		Transmission.
			Air gap.	Substance.	
		°C.	cms.	cms.	Per cent.
7·76 μ	Rhombic	18	15·70	7·50	47·8
7·79 μ	,,	96	15·75	8·00	50·8
7·79 μ	Prismatic	96	6·80	4·50	66·2
7·68 μ	Plastic	18	18·70	9·10	48·6
7·74 μ	Liquid	125	12·35	5·40	43·6
10·73 μ	Rhombic	18	33·50	11·85	35·4
10·80 μ	,,	96	33·30	11·20	33·6
10·80 μ	Prismatic	96	7·06	4·72	66·8
10·78 μ	Liquid	125	6·25	2·65	42·5
11·90 μ	Rhombic	18	6·40	0·90	14·1
11·97 μ	,,	96	6·00	1·10	18·34
11·97 μ	Prismatic	96	4·10	0·64	15·6
11·83 μ	Plastic	18	10·95	1·25	11·4
11·90 μ	Liquid	125	3·88	0·15	3·8

Results of Experiment.

The absorption and emission spectra are shown in the graphs, figs. 2 and 3. Curves A are for the spectra of rhombic sulphur at 18° C., the upper curves being for a thickness 0·67 mm. and the lower curve for a thickness of 6·0 mm. Curve B is the spectrum of prismatic sulphur at 96° C. and curve C that of rhombic sulphur at the same temperature. Curve D is the spectrum of liquid sulphur at 125° C. and curve E that of plastic sulphur at 18° C. The temperature shift between curves A and C is towards longer wave-lengths with increasing temperature, but is small and does not bear any obvious relation to the shift observed by Fukuda* in the visible, though it is in the same direction. The spectra of prismatic and rhombic sulphur at the same temperature are indistinguishable within the limits of experimental error. The spectrum of liquid sulphur is very similar to that of the solid phases, but the maxima of absorption appear to be broadened and slightly shifted towards shorter wave-lengths, in contra-distinction to the temperature shift. The spectrum of plastic sulphur, so far as it could be measured, is identical with that of liquid sulphur. The emission spectrum of liquid sulphur shows maxima which may be said to coincide within experimental error with the minima of the transmission spectra. The maximum deflection was only 3·5 mm. however. The values for the deflections obtained at the various maxima of the absorption spectra are given in

* Fukuda, 'Kyoto Coll. Science Mem.,' vol. 4, p. 351.

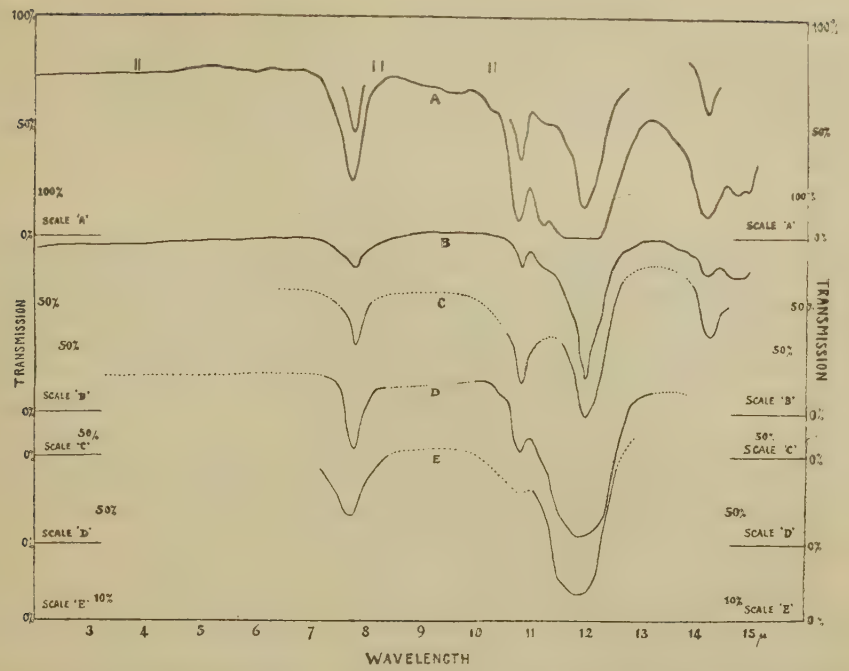


FIG. 2.—The Absorption Spectrum of Sulphur. Curve A, Rhombic sulphur at 18° C. Curve B, Prismatic sulphur at 96° C. Curve C, Rhombic sulphur at 96° C. Curve D, Liquid sulphur at 125° C. Curve E, Plastic sulphur at 18° C.

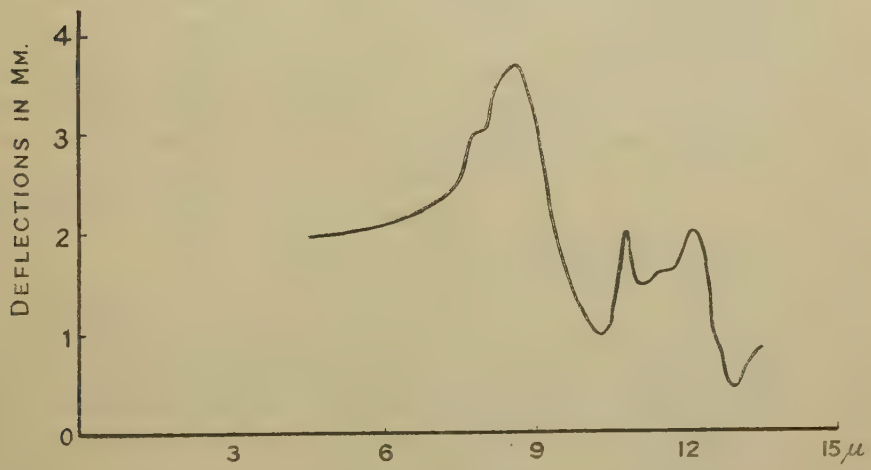


FIG. 3.—Emission Spectrum of Liquid Sulphur, *c.* 125° C.

Table II. The wave-length scale is accurate to about 0·01 μ and the ordinates which are proportional to the percentage transmission are accurate to about

$\pm 0{\cdot}5$ per cent. In the case of rhombic sulphur at 18° C. (curve A) the ordinates have the actual values of the percentage transmission, but for the other curves they are proportional to the transmission, the proportionality factor being as nearly unity as could be judged from the small specimens available.

Discussion of Results.

From the graphs it will be seen that the characteristic frequencies of vibration of sulphur are almost the same in all the phases—rhombic crystalline solid, prismatic crystalline solid, plastic or supercooled liquid, and molten liquid. This surprising result indicates that whatever complex is responsible for these vibrations, persists unchanged throughout the three phases. Further, the presence of active frequencies capable of affecting the passage of an electromagnetic wave points to the existence in this complex of an electric doublet, *i.e.*, the so-called "homopolar" linkage between the atoms is here akin to an heteropolar binding. Whether the atoms bear integral real charges or whether the doublet owes its being to some mutual polarisation of the atoms is not indicated. An important point to be noticed is the great depth of the chief absorption band at $11{\cdot}90\ \mu$. At the minimum of transmission only 14 per cent. of the incident energy is able to penetrate a slice of rhombic sulphur $0{\cdot}67$ mm. thick. Comparison of this with the fundamental absorption frequency of calcite at 14 μ* shows that the extinction coefficient of sulphur is about 1/10 that of calcite, which is a substance known to possess ions bearing large electrical charges (2 and 4 electrons). Sulphur atoms will be expected, therefore, to behave as if carrying an electric charge, and considered in pairs two neighbouring atoms must together form an electric doublet. It is not possible to estimate the moment of this doublet directly from the absorption spectrum, but an attempt is made in the following portion of the paper to calculate the mean value of the effective charge associated with the sulphur atom.

That the depth of an absorption band depends upon the damping of the oscillators concerned is shown by Victor Henri,† and an expression may be deduced which connects the extinction coefficient with the various characteristics of the oscillator. Thermal expansion shows that the forces called into play by any distortion of the crystal lattice must be asymmetric, but the phenomenon of superposition of wave disturbances, and the fact that the refractive index is independent of the intensity, leads to the view that the

* Reinkober, 'Z. f. Physik,' vol. 5, p. 192 (1921).

† Victor Henri, 'Photochimie.'

asymmetric term may be neglected for optical purposes. If a vibrating particle of mass m carry a charge ε, the equation of motion is

$$m\ddot{x} + r\dot{x} + f^2x - 4\pi\varepsilon D_1 = 0, \tag{1}$$

where

$$D_1 = D - A\varepsilon x,$$

and

$$D = D_0 e^{i(\omega t - qz)},$$

the direction of propagation of the impressed wave being along the z axis. $r\dot{x}$ is the frictional force giving rise to the damping and A the number of particles per cubic centimetre. The frequency ν of the impressed field due to the wave train is $\nu = \omega/2\pi$ and the natural frequency of the particles is

$$\nu_0 = 1/2\pi\sqrt{f/m}.$$

The solution of this equation leads to the well-known expressions

$$\mu^2 K = \frac{A\varepsilon^2 r\nu}{4\pi^2 m^2 (\nu_0^2 - \nu^2) + r^2\nu^2} \tag{2}$$

and

$$\mu^2 (1 - K^2) = 1 + \frac{4\pi m (\nu_0^2 - \nu^2) A\varepsilon^2}{4\pi^2 m^2 (\nu_0^2 - \nu^2)^2 + r^2\nu^2}, \tag{3}$$

where μ is the refractive index of the medium for waves of frequency ν and K an extinction coefficient such that the intensity of the wave is reduced in the ratio of $e^{-4\pi K}$ to 1 in travelling a distance λ/μ when $\lambda = c/\nu$ if c is the velocity of propagation in free space. If $I_0 e^{-\alpha z} = I_0$, where I_0 is the intensity of the incident wave and I the intensity of the mergent wave after traversing a thickness z of the medium,

$$\alpha = \frac{4\pi K\mu\nu}{c},$$

or substituting in (2)

$$\alpha = \frac{4\pi A\varepsilon^2 r\nu^2}{\mu c} \cdot \frac{1}{4\pi^2 m^2 (\nu_0^2 - \nu^2)^2 + r^2\nu^2}. \tag{4}$$

From this it is seen that α is small except in the region $\nu_0 = \nu$, and it becomes a maximum at the point $\nu = \nu_0$

$$\alpha_{max.} = \frac{4\pi A\varepsilon^2}{\mu c r}. \tag{5}$$

(4) may be written

$$\alpha = \frac{r/m}{4\pi^2\Delta^2 + r^2/m^2} \cdot \frac{4\pi A\varepsilon^2}{\mu c m}$$

if

$$\Delta = \nu_0 \left(\frac{\nu_0}{\nu} - \frac{\nu}{\nu_0}\right),$$

where Δ/ν_0 is the "mistuning" suggested by Barton in his treatment of sharpness of resonance,*

whence

$$\frac{\alpha}{\alpha_{max.}} = \frac{\mu_0}{\mu} \cdot \frac{r^2/m^2}{4\pi^2\Delta^2 + r^2/m^2},$$

whence

$$\frac{\mu\alpha}{\mu_0\alpha_{max.} - \mu\alpha} = \frac{r^2/m^2}{4\pi^2\Delta^2}.$$

Now if $\mu\alpha = \frac{1}{2}\mu_0\alpha_{max.}$, for some particular value of Δ

$$r/m = 2\pi\Delta. \tag{6}$$

From equation (3) $\mu_0^2 = 1/1 - K^2$, and as K is small in the infra-red spectrum $\mu_0 = 1$, and the condition for equation (6) to be true reduces to

$$\frac{\alpha}{\alpha_{max.}} = \frac{1}{2\mu}. \tag{7}$$

It is to be noticed that as a rule more than one natural period is possible to the medium and equation (3) becomes

$$\mu^2(1 - K^2) = 1 + \sum_{\nu_0}^{\nu_n} \frac{4\pi m(\nu_0^2 - \nu^2)A\varepsilon^2}{4\pi^2 m^2(\nu_0^2 - \nu^2)^2 + r^2\nu^2},$$

and the term under the summation sign is not zero for any value of ν. In this case μ_0 is not necessarily unity, and for values $\nu < \nu_0$, $\mu > \mu_0$, while for values of $\nu > \nu_0$, $\mu < \mu_0$.

It is not possible to obtain experimental information as to the magnitude of the μ, but an approximation to the truth will be made by putting $\mu_0 = \mu$. This approximation causes an error in the value found for r/m, because on one side of ν_0 the value of Δ^2 is greater than, and on the other side, less than, the true value. By taking the mean value for r/m as deduced from the magnitudes of Δ^2 on either side of the absorption maximum, the approximate result will not be seriously in error. Hence, if ν be chosen for two positions at which $\frac{\alpha}{\alpha_{max.}} = \frac{1}{2}$ and having regard to sign

$$r/m = \pi(\Delta_1 - \Delta_2). \tag{8}$$

The above treatment makes it possible to determine r from experimental measurements of transmission minima. The value of r so found may be inserted in equation (5) and the value of $\alpha_{max.}$ may then be estimated if A and ε can be assumed and if μ is put equal to unity—an approximation which is very near the truth for $\nu = \nu_0$, *i.e.*, at $\alpha_{max.}$

* Barton, 'Phil. Mag.,' vol. 26, p. 111 (1913).

Returning to equation (5), A may be written as $N\rho n/M$, where M is the molecular weight, n the number of vibrating particles per molecule, N is Avogadro's number, and ρ is the density of the substance, so that

$$\alpha_{max.} = \frac{N\rho n}{M} \frac{4\pi\varepsilon^2}{cr}. \tag{9}$$

This equation is not complete because it neglects any variation of $\alpha_{max.}$ with temperature, and the work of Reinkober* showed that, in the case of reflection from ammonium salts, lowering the temperature from 290° to 90° absolute increased the reflective power about $2\frac{1}{2}$ fold, and since

$$R = \frac{(\mu - 1)^2 + K^2}{(\mu + 1)^2 + K^2},$$

K, and therefore $\alpha_{max.}$ must vary in the same sense. Substitution of the appropriate numerical values in equation (9) gives a result for $\alpha_{max.}$ some 300 times greater than is found experimentally. It would therefore appear that only a small fraction of the particles which are theoretically capable of absorbing, do so in practice, and that this fraction decreases with increasing temperature.

Equation (9) may, however, be made to serve as a basis for estimation of the effective charge on the atom of sulphur, since if any two substances, say calcite and sulphur, be compared,

$$\alpha_{max.}\ \text{calcite} = \frac{N\rho_1 n_1}{M_1} \cdot \frac{4\pi\varepsilon_1^2}{cr_1}$$

and

$$\alpha_{max.}\ \text{sulphur} = \frac{N\rho_2 n_2}{M_2} \cdot \frac{4\pi\varepsilon_2^2}{cr_2},$$

and on division and rearrangement

$$\varepsilon_2 = \varepsilon_1 \sqrt{\frac{\rho_1 n_1 r_2 M_2}{\rho_2 n_2 r_1 M_1} \cdot \frac{\alpha_{max.}\ \text{sulphur}}{\alpha_{max.}\ \text{calcite}}}. \tag{10}$$

From the figures kindly provided by Prof. Schaefer for the absorption of calcite† $\alpha_{max.}$ occurs at 13·93 μ and has a value of 280. $\alpha = \alpha_{max.}/2$ occurs at 14·12 μ and 13·77 μ, whence from expression (8), $r/m = 3{\cdot}65 \times 10^{12}$, where $m = 16 \times 1{\cdot}64 \times 10^{-24}$ gm. is the mass of the oxygen atom, or

$$r = 0{\cdot}96 \times 10^{-10} \text{ dynes per unit velocity.}$$

From fig. 2, considering the most important absorption maximum for rhombic sulphur, $\alpha_{max.}$ occurs at 11·89 μ and has a value of 29, and $\alpha = \alpha_{max.}/2$ occurs

* Reinkober, 'Z. f. Physik,' vol. 3, p. 18 (1920).

† Schaefer, 'Z. f. Physik,' vol. 39, p. 648 (1926).

at 11·69 μ and 12·25 μ, whence $r/m = 7{\cdot}1 \times 10^{12}$, where m = mass of sulphur atom which is equal to $32 \times 1{\cdot}64 \times 10^{-24}$ gm., or

$$r = 3{\cdot}73 \times 10^{-10} \text{ dynes per cm./sec.}$$

It may here be noted that r/m is equal to $2/\tau$, where τ is the time of molecular relaxation, which has been shown by Rideal* to be of the order of the Lindemann melting-point frequency. For sulphur, ν Lindemann $= 3{\cdot}96 \times 10^{12}$, which agrees well with the value of $1/\tau = 3{\cdot}6 \times 10^{12}$ as found above.

The density of calcite $\rho_1 = 2{\cdot}72$ and of sulphur $\rho_2 = 2{\cdot}06$, while the molecular weight M_1 of calcite is 100 and M_2 may be taken for sulphur to be 64, corresponding to two atoms.

The vibrations of calcite in the infra-red have been analysed by Brester† from symmetry conditions, and it appears that the vibration frequency, found by Schaefer‡ at about 14 μ, is due to oscillation of the oxygen atoms, which bear a charge of two electrons. The number of vibrating atoms per molecule of $CaCO_3$ is then three, but owing to the configuration of the crystal one half of this number will be effective for any directed wave, so that $n_1 = 3/2$. n_2 may be taken to be 2, since all the sulphur atoms may be considered to be equally capable of response to an electromagnetic wave. Thus equation (10) gives

$$\varepsilon_2 = \varepsilon_1 \times 0{\cdot}505. \qquad (11)$$

Now ε_1 is the effective charge on the oxygen atoms in calcite, and this is less than the true charge owing to the effect of polarisation of the atoms, and Kornfeld§ has shown that only by a consideration of the polarisation as well as of the actual charges, can this frequency of vibration be calculated. Bragg and Chapman|| find from a consideration of the rhombohedral angle of calcite that the electric centre of the oxygen atom is distant 0·95 Å from the centre of the CO_3 group, whereas the X-ray internuclear distance between oxygen and carbon atoms is 1·25 Å. The effective charge on the oxygen atoms would thus appear to be 0·95/1·25 of the true charge, or $1{\cdot}52\,e$, where e is the charge on one electron. Substituting this value in (11) $\varepsilon_2 = 0{\cdot}77\ e$, so that the effective charge on the sulphur atom is found to be 0·77 electron, or nearly that which would be produced by complete transference of an electron from one atom to its neighbour.

This determination of the value of ε' the effective charge associated with each

* Rideal, 'Phil. Mag.,' vol. 40, p. 461 (1920).
† Brester, 'Z. f. Physik,' vol. 24, p. 324 (1924).
‡ Schaefer, 'Z. f. Physik,' vol. 39, p. 648 (1926).
§ Kornfeld, 'Z. f. Physik,' vol. 26, p. 205 (1924).
|| Bragg and Chapman, 'Roy. Soc. Proc.,' A, vol. 106, p. 369 (1924).

sulphur atom leads to the conclusion that sulphur behaves as if all the atoms were not precisely similar but that neighbouring atoms are so constituted as to appear to bear charges of opposite sign, and magnitude approximating to that of an electron. In other words, the outer electronic systems of each atom are so distorted that the linkage is approximating to that which exists when a single electron is completely transferred from one nucleus to the other, as is the case in heteropolar linking. The electric moment of the doublet composed of two atoms bearing charges $+ e'$ and $- e'$ at their centres of mass, these centres being separated by a distance x, is given by p, where $p = e'x$. p may also be written as equal to ex', where e is the actual charge displaced through a distance x'. In his Massachussetts lectures Born suggested that the difference in size of the atoms composing a crystal is an important factor in determining the order of symmetry of the lattice, great inequalities in size leading to low orders of symmetry. The order of symmetry of the rhombic system is not a maximum, and of the prismatic system it is still less; so that this fact may also be considered to point to a difference between the atoms composing the crystalline lattice of sulphur. From the foregoing remarks this difference is seen to be one of distortion of the electronic shells of each atom, resulting in neighbouring atoms bearing effective charges of opposite sign, pairs of atoms forming electrical doublets.

To obtain some idea of the manner in which these doublets are distributed in the crystal lattice, it is first of interest to review shortly the state of knowledge of the molecular weight of sulphur. The molecule has long been known to have a complicated structure, and it seems that in solution it may be large. When dissolved in phosphorus,* iodine,† and bromoform,‡ the molecular weight is approximately 260, and Popoff § showed that in many solvents the molecular weight was large, indeterminate, and increased with the concentration, being greater than S_8 in strong solution. Kellas‖ measured the surface tension of molten sulphur more accurately than previous observers, and came to the conclusion that the molecular weight of the pure liquid phase was probably S_8. It is to be noted, however, that he gives various other formulæ for deducing the molecular weight, and these lead to the values S_{18}, S_{14}, S_{12}, etc. Sulphur vapour is known to be almost entirely S_2 at temperatures

* Helff, 'Z. f. Phys. Chem.,' vol. 12, p. 196 (1893).
† Olivari, 'Atti. R. Acad. Lincei,' vol. 17, p. 512 (1908).
‡ Borgo and Amadori, 'Atti. R. Acad. Lincei,' vol. 18, p. 138 (1903).
§ Popoff, 'J. Russ. Phys. Chem. Soc.,' vol. 35, p. 642 (1903).
‖ Kellas, 'Chem. Trans.,' vol. 113, p. 903 (1918).

a little above the boiling point, but at lower temperatures the vapour becomes rapidly more dense, and Preuner and Schupp* suppose molecules S_8, S_6, S_4, S_3 and S_2 to exist. Graham,† examining the absorption spectrum of sulphur vapour in the visible and ultra-violet regions, came to the conclusion that only two types of molecule occur, which he assumes to be S_2 and S_8.

In the solid phase Bragg‡ found evidence for eight interpenetrating lattices, and Mark and Wigner§ by careful X-ray analysis show that rhombic sulphur has an ortho-rhombic diamond-type lattice of which the points are occupied by the centres of gravity of groups of 16 sulphur atoms, the unit crystallographic cell containing 128 atoms of sulphur. Bragg shows that in iron pyrites FeS_2|| the two sulphur atoms are very closely associated in pairs and the crystal is determined by the close-packed formation of these groups. He also points out that the distance apart of the sulphur nuclei is only 2·05 Å, a distance at which the repulsive forces must become very large, and it appears that the atoms are held together by mutual distortion of the electronic shells.

It appears that the most important groupings of sulphur atoms are S_2 and S_{16}, with a possibility of S_8 under certain conditions, and of these the S_2 is the most persistent configuration. It is tempting to suppose that the S_2 group is the one which is common to all phases of sulphur and even to FeS_2, and because the experimental work in the infra-red spectrum has shown that there is associated with pairs of sulphur atoms a considerable electric moment, it is clear that the aggregation of these groups into larger complexes would proceed upon normal lines dependent upon the electrostatic forces of attraction and repulsion between doublets. Such a structure would possess certain characteristics which enable calculations to be made of the thermal values of work of disruption of the crystal, but before proceeding farther some comparison may be made with other substances of more or less known structure. Taking the heats of sublimation and the compressibilities as the first criteria, it will be seen that as regards the former sulphur must be classed with those substances which form strongly polar ionic lattices, whereas regards compressibility sulphur would appear to stand midway between those elements and compounds which are thought to possess molecular lattices and those having ionic lattices.

* Preuner and Schupp, ' Z. f. Phys. Chem.,' vol. 68, p. 129 (1909).
† Graham, ' Roy. Soc. Proc.,' A, vol. 84, p. 311 (1910).
‡ Bragg, ' Roy. Soc. Proc.,' A, vol. 89, p. 575 (1914).
§ Mark and Wigner, ' Z. f. Phys. Chem.,' vol. 111, p. 398 (1924).
|| Bragg, ' Proc. Roy. Inst.,' vol. 24, p. 614 (1925).

Substance.	Heat of sublimation in cals. per gm. molecule.	Compressibilities $\times 10^{12}$.
Hydrogen halides	3,600→ 4,000	>30
Halogens	4,800→ 6,000	13
Sulphur S_2	30,000	13
Alkali halides	40,000→ 56,000	3→8

The heat of dissociation of sulphur is comparatively large, pointing to strong attractive forces between the atoms. That sulphur atoms may acquire effectively opposite polarity is also indicated by the great readiness with which, in the thionates and polysulphides, they may be linked into chains.

Returning to the suggestion that sulphur atoms should link in pairs to form units S_2 having an electric moment p, these units being built up into groups $(S_2)_8$ by virtue of the forces of attraction and repulsion between the doublets, it may be noted that a structure similar to this was suggested by Born and Kornfeld* for solid HCl, and they supposed the molecules of hydrochloric acid to be electrical dipoles and to be situated at the corners of a cube, pointing alternately towards and away from the centre. The S_2 unit has already been shown to be associated with an electrical moment, and it may be supposed to occupy a similar position in a rectangular group of 16 atoms, a suitable orthorhombic repetition of this group giving rise to the observed X-ray structure of sulphur, the centres of gravity of the groups $(S_2)_8$ forming a diamond type lattice. If the S_2 dipole be represented by an arrow →, the group is thus conceived on the basis of this hypothesis:

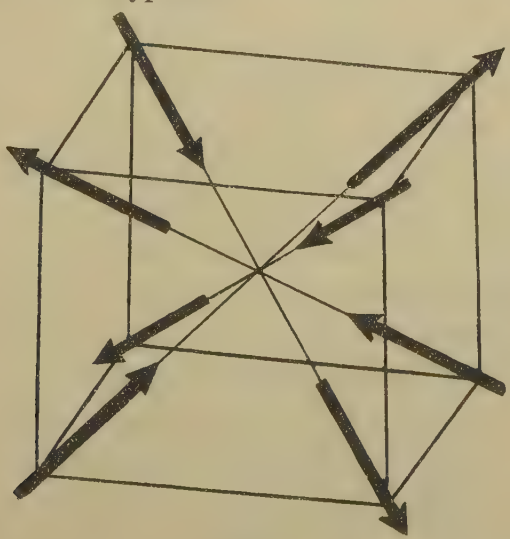


Fig. 4.

* Born and Kornfeld, 'Phys. Z.,' vol. 24, p. 121 (1923).

and it is possible to calculate the heat of vaporisation of sulphur by means of Kornfeld's formula involving the electric moment p. To do this it is necessary to make an approximation, replacing the actual ortho-rhombic lattice of sulphur by an equivalent cubic lattice of spacing equal to the average spacing of the actual structure. This approximation should give at least the right order of magnitude for the heat of sublimation.

Kornfeld's formula for the energy of such a " molecular lattice " is, assuming a repulsive force of potential $\rho\delta^{-n}$ between the groups, where δ is the period of repetition of the unit cell of side $\delta/2$

$$\phi = \frac{-21{\cdot}41}{\delta^3}\rho^2 + \frac{\beta}{\delta^n},$$

where ϕ is the lattice energy per molecule; from the equilibrium condition $d\phi/d\delta = 0$, β may be eliminated, giving

$$\phi = \frac{-21{\cdot}41}{\delta^3}\rho^2\left(1 - \frac{3}{n}\right),$$

therefore Q, the heat of sublimation per gram molecule, is

$$Q = N\phi\frac{1}{J} \text{ calories}$$

and for a cell containing 16 atoms of sulphur

$$\delta^3 = \frac{8M}{\rho N},$$

where M is the molecular weight of S_2, ρ the density and N is Avogadro's number, whence

$$Q = -\frac{N^2\rho}{8JM} \cdot 21{\cdot}41\ p^2\left(1 - \frac{3}{n}\right).$$

According to Lennard-Jones* n for sulphur is equal to 8.

Q may be taken as 30,000 cals. per gm. molecule (S_2), J as $4{\cdot}2 \times 10^7$ ergs per calorie, so that, giving the other constants the value used earlier in this paper,

$$p = 8{\cdot}0 \times 10^{-18} \text{ e.s.u.}$$

Making the assumption that the separation of the nuclei in the group S_2 is equal to that found by Bragg for the same group in iron pyrites then if $p = e'x$ where e' is the effective charge on the sulphur atoms,

$$\varepsilon' = 0{\cdot}82 \text{ electronic charge.}$$

This result agrees remarkably well with the value deduced from the experiments in the infra-red spectrum.

* Lennard-Jones, ' Roy. Soc. Proc.,' A, vol. 112, p. 230 (1926).

A further attempt to estimate p may be made from the experimental value for the heat of dissociation of sulphur. On the foregoing hypothesis, S_2 being of the type of a polar group having an electric moment, the dissociation may be considered to proceed in steps

$$S_2 \rightarrow S^+ + S^- - Q_2 \text{ calories},$$

$$S^- \rightarrow S + e - Q_3 \text{ calories},$$

$$S^+ + e \rightarrow S + Q_4 \text{ calories},$$

or if the total reaction

$$S_2 \rightarrow 2S - Q_1 \text{ calories},$$

then $Q_1 = Q_2 + Q_3 - Q_4$ calories.

The value of Q_4 is known to be that corresponding to an ionisation potential of 10·35 volts or $Q_4 = 248{,}000$ cals. per gram atom. Q_2 is not known, but Born's cycle applied to zinc sulphide gives a value for Q_5, where

$$S^{--} \rightarrow S + 2e - Q_5 \text{ calories}.$$

According to Born and Bormann* Q_5 corresponds to 2 volts, but Born and Gerlach† criticise this and point out that it should be much lower, while Franck and Jordan‡ give a different value for the spectroscopic data, and the derived value of Q_5 must accordingly be of the order of 0·2 volt or 5,000 calories per gram atom. Q_5 may be considered to be split into parts $Q_5 = Q_3 + Q_6$, where

$$S^{--} \rightarrow S^- + e - Q_6 \text{ calories}$$

and

$$S^- \rightarrow S + e - Q_3 \text{ calories}.$$

Since the S^{--} ion is stable in crystalline zinc sulphide, all the thermal values must be positive and small in the last two reactions, since their sum is small. As also Q_3 must be greater than Q_6, Q_6 may be neglected in comparison with Q_1, which is large, and $Q_3 \backsim Q_5$ and

$$Q_2 \backsim Q_1 + Q_4 - Q_5.$$

The most accurate value for Q_1 has been considered by Rideal and Norrish§ to be 103,000 calories per gram molecule, and taking this value

$$Q_2 \backsim 346{,}000 \text{ calories}.$$

The mechanism of dissociation is not known, and the calculation of the energy should involve the repulsive forces and the mutual polarisation of the atoms.

* Born and Bormann, 'Z. f. Phys.,' vol. 1, p. 250 (1920).

† Born and Gerlach, 'Z. f. Phys.,' vol. 5, p. 433 (1921).

‡ Franck and Jordan, "Structure der Materie," p. 281.

§ Rideal and Norrish, 'Trans. Chem. Soc.,' vol. 125, p. 2070 (1923).

But to a first approximation the energy may be considered as that work which would be done in dragging apart the components of an electric doublet p. If $p = \varepsilon x'$, where ε is the displaced charge and x' the distance between the centres of gravity of the electric charges distributed on each atom, then the work done in disruption of the doublet is $\frac{\varepsilon^2}{x'} = \frac{\varepsilon^3}{p}$ ergs or per gram molecule $\frac{N\varepsilon^3}{Jp}$ calories $= Q$. The previous estimates have shown p to be such that if $p = e'x = ex'$, e' is less than one electron. Therefore, the actual charge displaced may be assumed to be that due to the transfer of a single electron from one atom to another, so that, using the values of N and J employed earlier,

$$p = 4{\cdot}5 \times 10^{-18} \text{ e.s.u.},$$

or, as before, taking x to be 2·05 Å

$$\varepsilon' = 0{\cdot}46 \text{ electronic charge.}$$

Considering the nature of the simplifications made in this calculation, the result compares very favourably with previous estimates.

Conclusion.

The results of the infra-red examination of various forms of sulphur has shown the existence of electromagnetically active vibrations, and it is inferred that electric doublets are present in the sulphur complex. The doublets are believed to be units S_2, and these are built into groups $(S_2)_8$, corresponding to the groups of mass 512, the centres of gravity of which have been shown by Mark and Wigner to occupy the points of the diamond-type space lattice of rhombic sulphur. The frequencies ascribed to S_2 persist throughout the rhombic, prismatic, and liquid phases. Calculation of the electric moment of the S_2 group from the heat of vaporisation and the heat of dissociation of sulphur gives results agreeing well with the estimates made from considerations of the infra-red absorption spectrum. The linkage between atoms in the element sulphur is thus believed to be pseudo-heteropolar. In this connection it may be remarked that an examination of the —S—S— linkage in the thiosulphates would be of interest. The following table gives the results obtained for the electric moment p associated with the unit S_2 and also for e', the effective charge on the atoms; $p = e'x$, where x is the distance apart of the centres of mass of the atoms, and has been taken to be 2·05 Å.

Method.	Data.	p in e.s.u. $\times 10^{-18}$.	ϵ' in fractions of $4 \cdot 774 \times 10^{-10}$ e.s.u.
Formula for extinction coefficient	Absorption spectra	7·5	0·77
Kornfeld's formula	Heat of vaporisation	8·0	0·82
Approximate heat of dissociation	Heat of dissociation.....	4·5	0·46

We should like to take this opportuuity of thanking Prof. T. M. Lowry, F.R.S., and also Prof. A. Hutchinson, F.R.S., for the loan of instruments, and Prof. Clemens Schaefer, now of Breslau, for his kindly interest in the work. Our thanks are also due to the Council of the Royal Society and to Messrs. Brunner Mond for grants in aid of apparatus.

The Total Ionisation due to the Absorption in Air of Slow Cathode Rays.

By J. F. Lehmann, M.Sc., sometime 1851 Exhibition Scholar from the University of Alberta, and T. H. Osgood, M.Sc., sometime Carnegie Scholar from the University of St. Andrews.

(Communicated by Sir Ernest Rutherford, P.R.S.—Received May 25, 1927.)

The object of the experiments here described is to measure the average ionisation produced by the absorption in air of an electron with definite initial energy. From this the average expenditure of energy associated with the formation of a pair of ions can be estimated. The initial energies considered ranged between 200 and 1000 electron-volts.

Experiments on ionisation by electronic impact have generally been concerned, either with a determination of the ionisation potential of the gas, or with the ionisation per unit path due to an electron having a definite energy. The ionisation potential has been measured by determining the minimum energy a stream of electrons must have in order to ionise, even occasionally, a normal atom. It represents the energy expended by the ionising electron if no kinetic energy be transferred to either of the ions formed. If at the impact an atomic electron were ejected with appreciable kinetic energy, the energy expended by

the ionising electron would be correspondingly increased above the ionisation potential. Also electrons may dissipate their energy by processes other than ionisation, notably by excitation and by dissociation of diatomic molecules. For these reasons the average expenditure of electronic energy per pair of ions should exceed the ionisation potential. The excess of this average energy would indicate the extent to which processes other than ionisation contribute to the dissipation of the initial kinetic energy of the electrons. The purpose of the present experiments is to obtain further information on this phase of the ionisation problem.

Some attempts have been made to estimate the total ionisation and the average energy expended per ion pair from the data available concerning the ionisation per unit path and the range of the electrons. If i represents the ionisation per unit path due to an electron which has travelled a distance x cms., from an origin at which it had a kinetic energy equivalent to V volts, R the range of the electron, then I_ν, the total ionisation due to the complete absorption of this electron is given by

$$I_\nu = \int_0^R i\,dx.$$

This integral has been considered by Lenard* and by Kulenkampff.† Their values for I_ν are very discordant and neither is in agreement with the experimental values we have found. As Kulenkampff points out, the discrepancies are associated with the determination of the ionisation per unit path as a function of the distance the electron has travelled. It is usual to express the ionisation per unit path,‡ found experimentally, as a function of the energy of the electron. It remains to correlate the energy of the electron with the distance it has travelled. On this point there is little information available, for the range of electron energies under consideration, and an experimental determination of the total ionisation would appear to have many advantages.

Experimental method.

The method used in the present experiments is illustrated in fig. 1. Electrons were ejected from the hot tungsten filament F, charged to a definite negative potential with the respect to the face of the ionisation chamber B. A small fraction of the electrons passed through the capillary C into the ionisation

* Lenard, ' Quantitative über Kathodenstrahlen,' p. 143, *et seq.*

† Kulenkampff, ' Ann. d. Physik,' vol. 79, p. 97 (1926).

‡ The experimental data concerning ionisation per unit path is critically examined by K. T. Compton and Van Voorhis, ' Phys. Rev.,' vol. 27, p. 274 (1926).

chamber. These electrons could either be collected in the Faraday cylinder D or, when this cylinder was removed, allowed to enter the body of the ionisation chamber and ionise the gas therein. The ions were collected on one of the

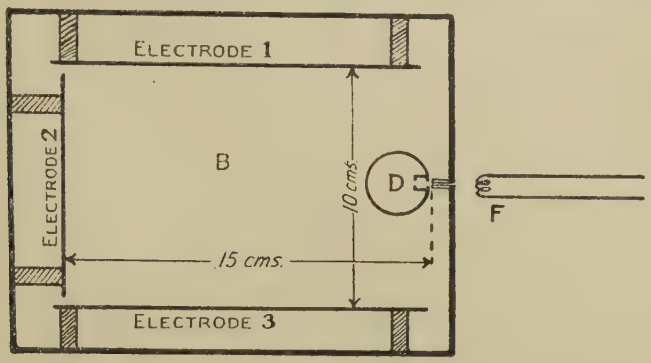


FIG. 1.

brass plates 1, 2, or 3 (fig. 1). The ionisation current divided by the incoming electron current, as measured by the Faraday cylinder, gave the average ionisation per electron.

Apparatus.

Fig. 2 is a sketch of a vertical section through the filament and ionisation chambers. A horizontal section is indicated in fig. 1, which shows the distribution of the electrodes used to collect the ionisation currents. A more detailed sketch of the filament chamber and of the Faraday cylinder, together with a diagram of the electrical connections is shown in fig. 3.

The filament F (fig. 2) was a tungsten wire 4 cms. long and 0·008 cm. in diameter, wound into a helix. It was heated by a current of from 0·8 to 1·2 amperes. Electrons were accelerated by a potential difference applied between the filament and the anode A. The electron current entering the ionisation chamber was measured by the rate of charge of a quadrant electrometer connected to the Faraday cylinder D. The Faraday cylinder was then raised by means of the windlass W and suitable guides insulated from the walls of the chamber by ebonite supports. The positive ions produced by the absorption of the electrons were collected on one of the electrodes 1, 2 or 3 (fig. 1) by applying a suitable potential difference between this electrode and the remaining electrodes and the walls of the chamber. Occasionally the negative ions were measured by reversing the collecting potential difference, but these measurements required to be corrected in order to allow for the charge carried by the ionising electrons. Provided that the electron beam was completely absorbed by the gas in the

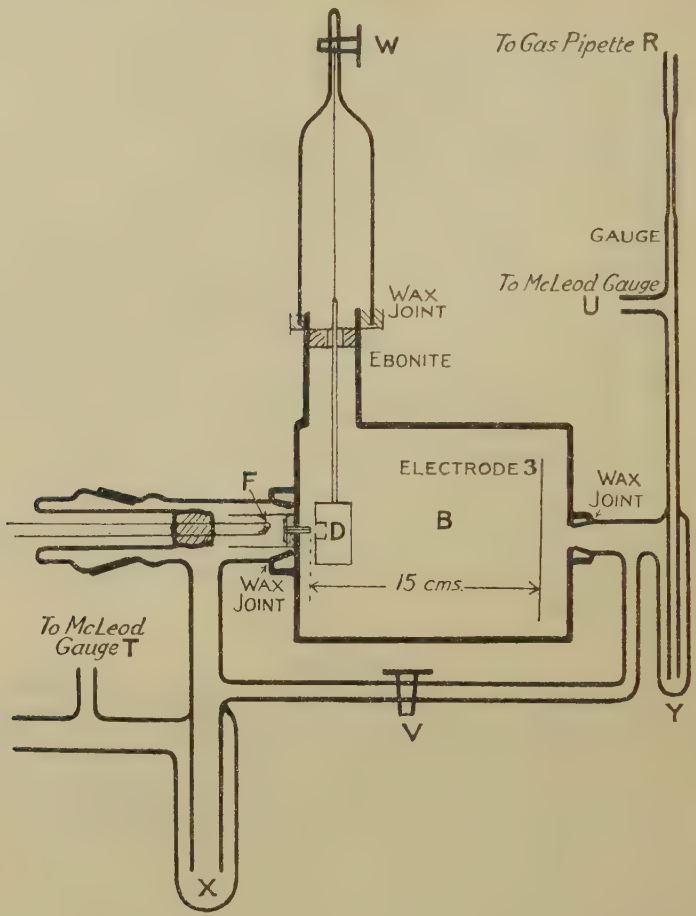


Fig. 2.

chamber, the difference between the negative and positive ionisation currents was found to be equal to the incoming electron current within the limits of experimental error.

The quadrant electrometer used to measure the electron and ionisation currents was connected in parallel with a condenser of variable capacity. The rate of charge of the condenser-electrometer system was measured in the usual way. Currents ranging from 10^{-12} to 10^{-8} amperes could be measured with an accuracy of about 1 per cent. by adjusting the capacity of the condenser. Generally it was convenient to work with currents of the order of 10^{-9} amperes and the filament heating current was adjusted accordingly.

Gas leaking through the capillary C was removed from the filament chamber

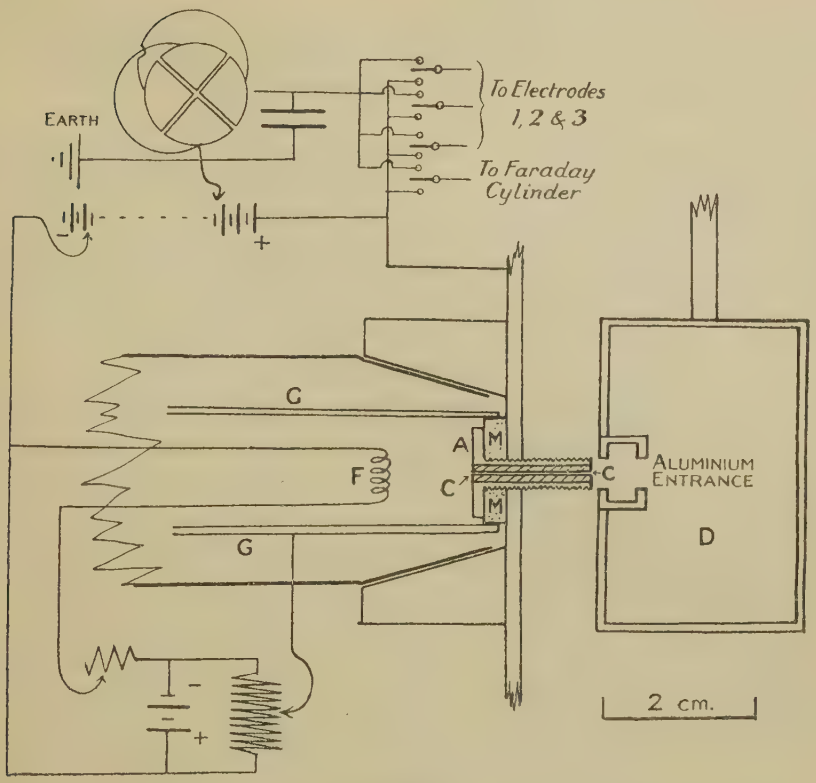


FIG. 3.

by high speed diffusion pumps and the pressure in the ionisation chamber was maintained by a capillary leak from a gas pipette. The highest pressure maintained in the ionisation chamber was about 2 mm. Hg, and under these conditions the pressure in the filament chamber was about 0·001 mm. Hg.

To obtain trustworthy results the following conditions had to be fulfilled :—

(1) The incoming electron beam must be homogeneous, *i.e.*, on entering the ionisation chamber each electron had the full energy corresponding to the potential difference through which it had been accelerated in the filament chamber.

(2) The Faraday cylinder must be designed to prevent the escape of any electrons, either primary incoming electrons or secondary electrons ejected by impact of the primaries on the wall of the cylinder.

(3) The potential difference used to collect the ions during the measurements of the ionisation currents must be sufficient to prevent any recombination and yet insufficient to cause ionisation by collision.

(4) In addition the possibility that radiation excited in the ionisation chamber might eject photo-electrons from the surface of the electrode used to collect the ionisation currents had to be carefully tested.

1. *The attainment of Homogeneous Electron Beams.*—It was discovered early in the course of this investigation that, in general, electrons which have passed through an aperture in an anode are by no means homogeneous,* the degree of homogeneity depending to a great extent on the dimensions of the aperture. The measurements relating to circular holes in copper sheets have been verified by Compton and Van Voorhis† and by Breit and Whiddington.‡ Those relating to copper capillaries have been questioned by Lawrence.§

Lawrence repeated our experiments using Faraday cylinders of variable depth to collect his electrons. He found that the apparent homogeneity of the beam improved as the depth of the cylinder increased. This effect was due to the escape of electrons from the cylinder under the influence of the retarding potentials used to analyse the incoming electron stream. The escape of electrons decreased with increasing depth of cylinder. The apparent homogeneity found by Lawrence became about constant and independent of the dimensions of the cylinder for deep cylinders. This constant value showed a homogeneity slightly better than the value we published. It is probable that our results may be subject to a correction of the order of 5 per cent. Even with his longest cylinders Lawrence found a marked inhomogeneity, and it seems reasonable to suppose that this is real and not associated with the escape of electrons from the Faraday cylinder.

Since the publication of these results considerable progress has been made in the attainment of a homogeneous electron stream. The homogeneity of the beams was measured by evacuating the ionisation chamber and applying retarding potentials between the Faraday cylinder D and the capillary C (figs. 1 and 2). The results are shown in fig. 4 where the current measured by the Faraday cylinder is plotted against the applied retarding potential. It is at once evident that a high degree of homogeneity has been attained. The assumption of perfect homogeneity cannot introduce an error of more than about 1 per cent.

This homogeneity was obtained by means of the arrangement indicated in fig. 3. Previous experiments had shown that the lack of homogeneity was

* Lehmann and Osgood, ' Proc. Camb. Phil. Soc.,' vol. 22, p. 731 (1925).

† Compton and Van Voorhis, *loc. cit.*

‡ Breit and Whiddington, ' Proc. Leeds Phil. Soc.' (1926).

§ Lawrence, ' Bull. Amer. Acad. Sci.' (Dec., 1925).

probably due to the presence of secondary electrons ejected from the anode by impact of the primary electrons. To minimise the emission of secondaries the

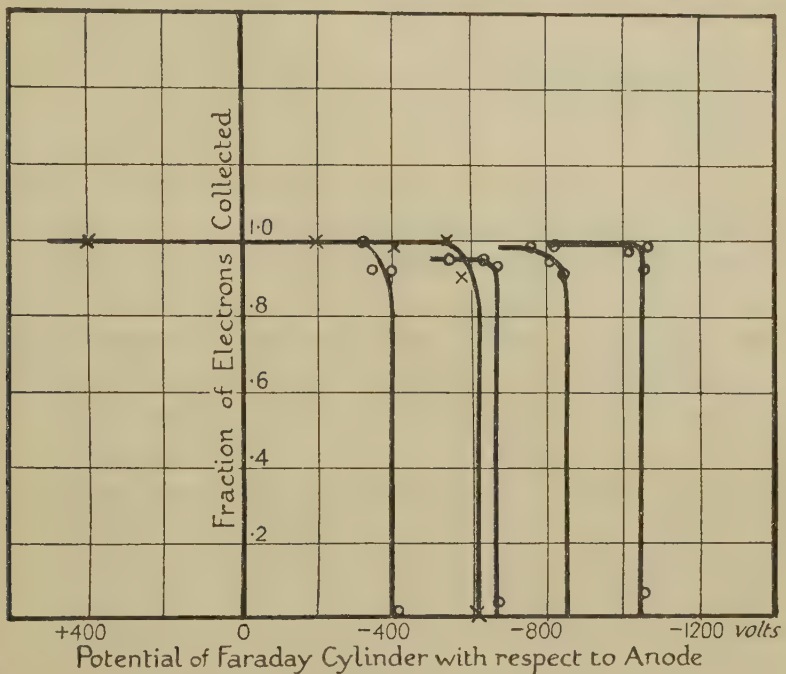


Fig. 4.—Homogeneity of Electron Beams.

electrons were focussed to pass through the capillary exactly parallel to its axis, thus reducing to a minimum impacts with the walls of the capillary. It was also found advisable to use as capillary a graphite rod with a hole 0·03 cm. in diameter drilled down its axis. An aluminium capillary gave considerably more homogeneous beams than did a copper one and graphite was slighter better than aluminium.

The electrons were focussed by means of a cylindrical shield G (fig. 3) kept at the potential of the filament which it surrounded. This shield was mounted coaxially with the anode capillary and a slight mal-adjustment resulted in an inhomogeneous beam. To assure that the shield maintained its proper position it was mounted on a marble collar M. The stem of the anode A, in which was mounted the graphite capillary C, passed through a hole in the centre of this collar and screwed into the face of the ionisation chamber. With this arrangement the anode and shield could be removed and replaced in exactly the same position.

It was also essential that the filament should be centred opposite the entrance

to the capillary C. To facilitate its removal and replacement, the filament was mounted on a ground glass joint. During the glass-blowing care was taken to mount the filament as nearly as possible on the axis of the shield G. The final adjustment was made by rotating the ground joint until the electron current entering the ionisation chamber was a maximum.

The effectiveness of the focussing arrangement could be judged by observing a black spot on the anode which was easily visible after a run of about half-an-hour and increased in intensity with time. This black spot was no doubt due to negatively charged particles of tungsten sputtered from the filament. The radius of the spot was of the order of two or three millimetres. It seems reasonable to suppose that the electrons would be more sharply focussed than the tungsten particles, and therefore strike the anode at the centre of the black spot. Homogeneous electron beams were obtained only when the centre of the spot coincided with the entrance to the capillary. If the centre of the spot were displaced by as much as 2 mm. no electrons passed through the capillary.

Another factor which affected the homogeneity of the electrons entering the ionisation chamber was the distance between the filament and the face of the anode. For a given distance between these electrodes there was a certain range of accelerating potentials within which homogeneous beams were obtained. To work with accelerating potentials below this range it was necessary to move the filament closer to the anode; for higher accelerating potentials it was advisable to increase the distance between filament and anode. With the filament 0·5 cm. from the anode homogeneous beams were obtained when accelerating potentials ranging between 200 and 350 volts were used. At 1·0 cm. the range of accelerating potentials was from 300 to 600 volts, and at 1·5 cm., from 400 to 1,000 volts. These results may be explained by assuming that homogeneous electron beams were obtained when there was a uniform potential gradient in the neighbourhood of the entrance to the anode capillary. This potential gradient would be determined by the accelerating potential difference, the distance between filament and anode, and the space charge. The results suggest that, for a given distance between filament and anode, there was a definite range of accelerating potentials within which an adequately uniform potential gradient was obtained.

When the filament was close to the anode the homogeneity of the electron beam was very sensitive to slight changes in the position of the filament. At a distance of 0·5 cm., a 2 degree rotation of the ground joint supporting the filament produced a marked decrease in the homogeneity of the beam entering the ionisation chamber. A 5 degree rotation reduced the electron current enter-

ing the ionisation chamber to about 5 per cent. of its maximum value. This effect was still more pronounced when the filament was 0·2 cm. from the anode. At this distance the homogeneity was always poor and was subject to wide fluctuations. The shortest distance between filament and anode at which it was possible to carry through a prolonged series of ionisation measurements was 0·5 cm. The lowest potential which could be used at this distance was about 200 volts.

It is evident that the earth's magnetic field must affect the passage of the electrons through the anode capillary. In the vertical component a 300-volt electron would have a radius of curvature of about 1 metre. This would be ample to deflect many of the electrons against the walls of the anode capillary. To neutralise the earth's field the apparatus was surrounded by a pair of Helmholtz coils 65 cm. in diameter mounted with their axis parallel to the total force as determined by a dip needle. Before the apparatus was assembled the current through these coils was adjusted so make the period of a small oscillation magnetometer placed at their centre a maximum. This current was passed through the coils whenever ionisation measurements were taken. The apparatus was assembled with the anode capillary at the centre of the Helmholtz coils.

The homogeneity of the electron beams entering the ionisation chamber was always measured with the apparatus evacuated, to a pressure less than 0·0001 mm. Hg. During ionisation measurements the pressure in the ionisation chamber was at times as high as 2 mm. Hg. The question arises as to whether this change in conditions affected the homogeneity of the electron beams.

In Table I it will be observed that, with a constant filament heating current, the electron current entering the filament chamber decreased as the pressure n that chamber increased. This effect might be due to one of two causes. Either, during their passage through the anode capillary a considerable fraction of the electrons experienced collisions with gas molecules and were deflected out of the beam, or the emissivity of the filament decreased due to the increasing pressure in the filament chamber, which must accompany an increase in pressure in the ionisation chamber. Such a decrease in emissivity might be due to chemical changes at the surface of the hot tungsten, or to a slight decrease in the temperature of the filament as the conductivity of the surrounding gas increased with increasing pressure.

The pressure measured by the McLeod gauge T of fig. 2 ranged up to 0·001 mm. Hg as the pressure in the ionisation chamber ranged up to 2 mm. Hg.

Table I.—Ionisation by 410 volt electrons.

Electron current.	Positive ion current.	Positive ions per electron.	Pressure.
$1{\cdot}04\times10^{-9}$ amps.	$1{\cdot}7\times10^{-9}$ amps.		
1·83	2·0	2·1	0·02 mm. Hg
	2·1		
1·58	2·9		
1·59	2·7	4·8	0·055
	2·8		
1·35	2·8	7·5	0·10
1·37	2·6		
1·24	2·2	8·8*	0·14
1·26	2·2		
1·21	2·0	9·3*	0·16
1·20	1·8		
1·17	1·6	8·6*	0·19
1·18	1·4		
1·14	1·3	9·3*	0·22
1·14	1·3		
1·15	1·3	8·9*	0·23
1·13	1·2		

No. of positive ions per electron for complete absorption. (Mean of *) = 9·0.

To determine whether the decrease in the filament current was due to a pressure gradient in the anode capillary or to the increase in pressure in the filament chamber, a subsidiary run was made with the by-pass joining the ionisation and filament chambers open. The pressure in both chambers was raised from the limits of measurements of the McLeod gauges up to 0·005 mm. Hg. A similar decrease in electron current entering the ionisation chamber was observed as had previously been observed with the by-pass closed and the pressures as quoted in the Table I. We conclude that the decrease in electron current was due to a decrease in filament emission rather than to impacts of electrons with gas molecules. It seems improbable that such a decrease in emission affected the homogeneity of the electron beam. The emission was varied, *in vacuo*, by changing the filament heating current and no change was noted in the homogeneity of the electron beams.

Further evidence that the homogeneity of the electron beam is unaffected by the pressure in the ionisation chamber may be deduced from the experimental results shown in fig. 5. The curves show that, provided the pressure is adequate to absorb the electron stream, the average ionisation per electron is independent

of the pressure in the ionisation chamber. Were the energy of the electron beam to decrease with increasing pressure we should expect a corresponding decrease in the average ionisation per electron. No such effect was observed.

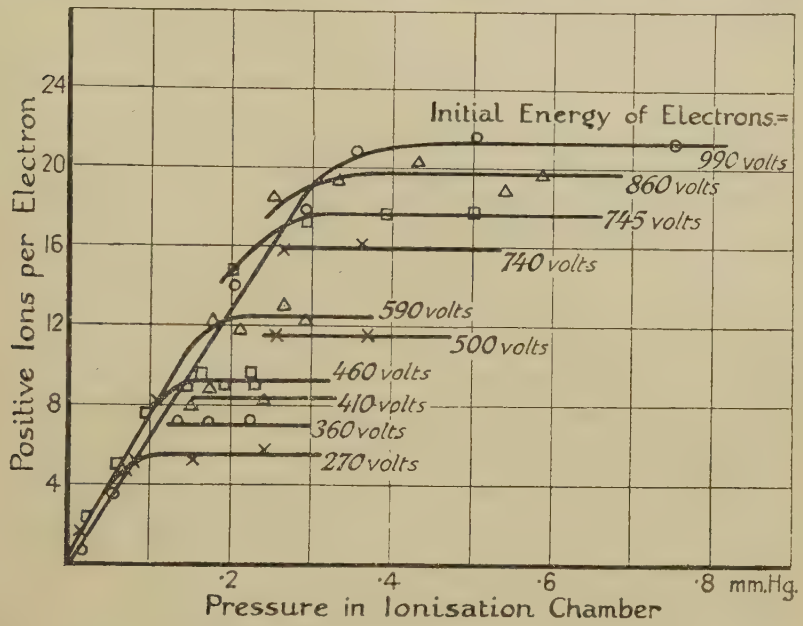


Fig. 5.—Ionisation in Helium.

2. *The Faraday Cylinder.*—The Faraday clyinder D, of fig. 3, was designed to prevent the escape of any electrons, either primaries entering it and reflected from its walls, or secondaries ejected by the impact of the primaries. The entrance to this cylinder was a small aluminium cylinder 6 mm. in diameter and 6 mm. long with two holes 4 mm. in diameter at each end. The electrons passed through these two holes into the body of the Faraday cylinder. The purpose of this entrance was to shield the interior of the Faraday cylinder from a stray field when retarding potentials were applied to analyse the electron beams. The presence of such stray field would greatly facilitate escape of the electrons as it would tend to draw them to the anode. The opening into the Faraday cylinder subtended at the opposite wall a solid angle of 0·02 radians so that the chance escape of electrons must have been small. That the arrangement was effective is shown by the curves of fig. 4 which show no indication of any escape of electrons.

When in position to collect the incoming electrons the centre of the Faraday cylinder entrance was placed 1 mm. from the end of the anode capillary. To

escape entering the cylinder an electron would have to be deflected through an angle of at least 60 degrees in this short distance. It seems highly probable that the measurements of the incoming electron current cannot be in error by more than 1 per cent.

3. *The Potential Difference used to Collect the Ions.*—Subsidiary experiments were made to determine the potential difference between the collecting electrode and the walls of the ionisation chamber, which was just adequate to produce saturated ionisation currents. Using electron beams of various initial energies absorbed in several different gases, this potential difference was found to be about 10 volts. Between 10 and 80 volts the ionisation current remained constant. In all the measurements quoted in this and the following paper the potential difference used to collect the ions was 20 volts.

4. *Photo-électron emission from the collecting electrode.*—In addition to ionising the gas in the ionisation chamber the electron stream must " excite " some of the gas atoms. These when regaining their normal state would emit radiation, and it was thought that some of this radiation might penetrate to the electrode used in measuring the ionisation current and eject photo-electrons from its surface. To test the point the ionisation current was measured, first with only one of the electrodes, 1, 2, or 3 (fig. 1), connected to the electrometer, the remainder being connected to the walls of the chamber ; then with two of the electrodes connected to the electrometer ; and finally with all three electrodes so connected. In this way the area used to collect the ions was varied from 160 sq. cm. to 400 sq. cm. Were there an appreciable emission of photo-electrons the positive ion current would appear to increase with increasing area of electrode. No such effect was observed, and it was concluded that the ejection of photo-electrons did not affect the measurements of the ionisation current by as much as 1 per cent.

Experimental Procedure.

Before any ionisation measurements were made the homogeneity of the electron beam to be used was tested *in vacuo*. If the homogeneity was satisfactory the by-pass, V (fig. 2), was closed and gas was introduced into the ionisation chamber from the gas pipette, R. The pressure in that pipette was adjusted to give the required equilibrium pressure in the ionisation chamber.

When the equilibrium pressure was established in the ionisation chamber, the electron stream entering the chamber was measured by collecting it in the Faraday cylinder. The cylinder was then raised and the positive ionisation current was measured. Generally two measurements of the ionisation current

were taken before the Faraday cylinder was lowered to check the electron current. By taking alternate readings of the ionisation and electron currents errors due to a fluctuating filament emission were minimised. The ratio of the ionisation current to the electron current gave the average number of positive ions produced by each electron.

Results.

The experimental results are given in figs. 5, 6 and 7. Fig. 5 shows the average ionisation per electron, as a function of the pressure in the ionisation chamber. for the different initial energies of the electron beams. Fig. 6 shows the ionisation per electron, for complete absorption of the electrons, as a function

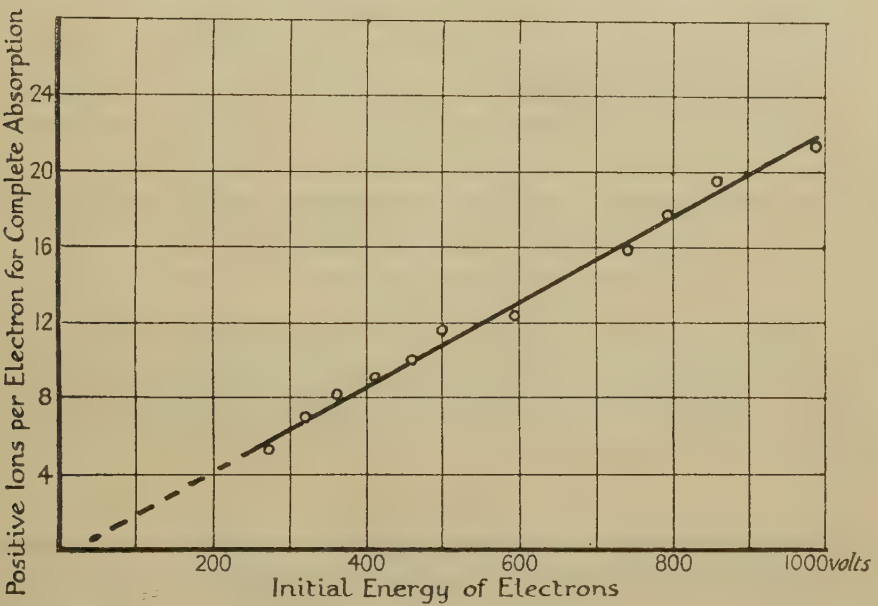


Fig. 6.

of the initial energy. A typical set of measurements is tabulated in Table I ; the mean of readings " * " gives the average number of positive ions per electron for complete absorption of the electrons. The tabulated results are shown in the 410-volt curve, fig. 5.

From fig. 5 it is evident that, with an electron beam of given initial energy, the ionisation per electron is proportional to the pressure at low pressures. For pressures greater than a certain " critical pressure" the ionisation per electron is constant and independent of pressure. The critical pressure is just adequate to absorb the electrons before they reach the walls of the ionisation

chamber. Assuming that the electrons are subject to but slight lateral diffusion, their range at the critical pressure must be the length of the ionisa-

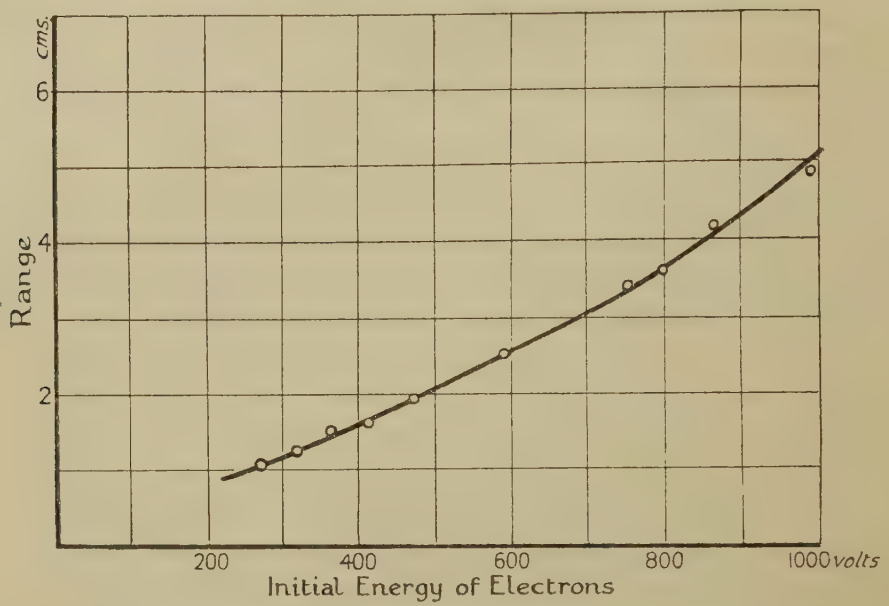


Fig. 7.—Range of Electrons in Air (pressure = 1 mm. Hg.).

tion chamber, viz., 15 cm. From these data the range at any standard pressure may be estimated by assuming that the range is inversely proportional to the pressure. In fig. 7 the estimated range for a pressure of 1 mm. Hg is plotted against the initial energy of the electrons.

Discussion.

We may estimate the average energy expended in the formation of a pair of ions by dividing the number of positive ions per electron, for complete absorption, into the initial energy of the electrons. From fig. 6 it is evident that in air this average expenditure of energy per ion pair is 45 electron-volts.

Fig. 6 shows that the ionisation per electron I may be expressed as—

$$I = 0{\cdot}0225\ (V - 17)$$

where V is the initial energy of the electron in volts and 17 is the ionisation potential of air.

Using a different experimental method Johnson* measured the total ionisation per electron for initial electron energies below 200 volts. He found

$$I = 0{\cdot}0275\ (V - 11).$$

* Johnson, 'Phys. Rev.,' vol. 10, p. 609 (1917).

The agreement is probably as good as could be expected in view of the fact that Johnson did not have the facilities of modern high-speed diffusion pumps, so that the pressure in his filament chamber had to be the same as that in his ionisation chamber.

The present results do not agree with those of Miss G. A. Anslow,* who used a method similar to the present one and covered the same range of electron energies. Miss Anslow did not find a linear relationship between I and V. For slow electrons, with initial energies below 500 volts, her values of the ionisation per electron are about half those we find. Between 500 and 1,000 volts her values for I rapidly increase and at 1,000 volts they are about twice as large as the values we find, viz., at 1,000 volts I (Miss Anslow) = 43 ions per electron, and I (Lehmann and Osgood) = 22 ions per electron. At 890 volts Miss Anslow found 24·1 ions per electron, at 952 volts 36·7 ions per electron. Thus an energy expenditure of 62 electron-volts produced 12·5 ions, representing an average expenditure of 5 electron-volts per ion pair. This is difficult to reconcile with the fact that the ionisation potential of air is 17 volts. Miss Anslow did not state the homogeneity of her electron beams, and it is possible that some of the discrepancies may be associated with the use of inhomogeneous beams.

These experiments have been continued and ionisation measurements have been made in several other gases. Further discussion of the measurements in air is postponed to a later paper wherein they may be correlated with the corresponding measurements in other gases.

Summary.

Homogeneous beams of slow cathode rays were introduced into an ionisation chamber. The kinetic energy of the electrons was absorbed by the air in the chamber and the resulting ionisation was measured. The ratio of ionisation current to electron current gave the average number of ions produced by the absorption of each electron. The electrons could be given any desired initial energy within the range 200 to 1,000 electron-volts.

The average ionisation per electron, for complete absorption, was found to be proportional to the initial energy of the electrons.

The average energy expended in the formation of a pair of ions was found to be 45 electron-volts in air.

The range of the electron beams was measured.

* Miss G. A. Anslow, ' Phys. Rev.,' vol. 25, p. 484 (1925).

The writers are pleased to acknowledge their indebtedness to Sir Ernest Rutherford, and express their thanks for his continued interest and advice. We also wish to thank Dr. Chadwick and Dr. R. J. Clark, whose assistance and suggestions on points of technique have been very helpful.

The Absorption of Slow Cathode Rays in Various Gases.

By J. F. Lehmann, M.Sc., sometime 1851 Exhibition Scholar from the University of Alberta.

(Communicated by Sir Ernest Rutherford, P.R.S.—Received May 25, 1927.)

Recently a method has been developed whereby the total ionisation due to the absorption of slow cathode rays may be directly measured. The application of this method to measurements of the total ionisation in air was the subject of a previous paper. The present paper concerns its application to helium, argon, hydrogen, nitrogen, and carbon dioxide.

A beam of electrons, all having the same initial energy, was introduced into an ionisation chamber containing gas at a given pressure. The incoming electron current and the ionisation current due to the passage of the electrons through the chamber were measured. The ratio $\frac{\text{ionisation current}}{\text{electron current}}$ determined the average ionisation per electron.

A detailed description of apparatus and procedure has been given previously.*

There follows a statement of the purity of gases used in the ionisation measurements.

Helium.—The supply of helium available consisted of a mixture 80 per cent. helium and 20 per cent. nitrogen stored at atmospheric pressure over water. It was purified by passing slowly through charcoal immersed in liquid air into the gas pipette supplying the ionisation chamber. On its passage to the ionisation chamber the gas was again passed through charcoal immersed in liquid air. The spectrum excited in a Geisler tube attached to the ionisation chamber was examined in a Hilger constant deviation spectrometer ; only slight traces of hydrogen and mercury vapour were detected.

Hydrogen.—Commercial hydrogen was stored over charcoal immersed in liquid air and the unabsorbed residue was used in the ionisation measurements. Spectroscopic examination showed but slight traces of mercury vapour.

* Lehmann and Osgood, *supra*, p. 609.

Argon.—Two samples of argon were used.

Sample 1 was purified from commercial argon, a mixture 80 per cent. argon and 20 per cent. nitrogen. The mixture was circulated in an iron furnace, containing a charge of calcium and magnesium at 900° C., until a spectroscopic examination showed no band spectrum. The gas was then pumped over hot copper oxide and through drying tubes into the gas pipette supplying the ionisation chamber. Spectroscopic examination of the gas in the ionisation chamber showed but slight traces of hydrogen and mercury vapour.

Sample 2 was taken from a cylinder of argon supplied by the British Oxygen Company and stated to contain from $\frac{1}{2}$ to 1 per cent. nitrogen.

The ionisation measurements of the two samples agreed to within the limits of experimental error.

Nitrogen.—Commercial nitrogen was used without any further purification.

Carbon Dioxide.—Commercial carbon dioxide was used. The liquid air on the traps leading to the ionisation chamber and to the diffusion pumps was replaced by solid carbon dioxide moistened with ether.

Results.

The results are shown in figs. 1 to 4, where the average number of ions per electron, for given initial electron energies, is plotted against the pressure in

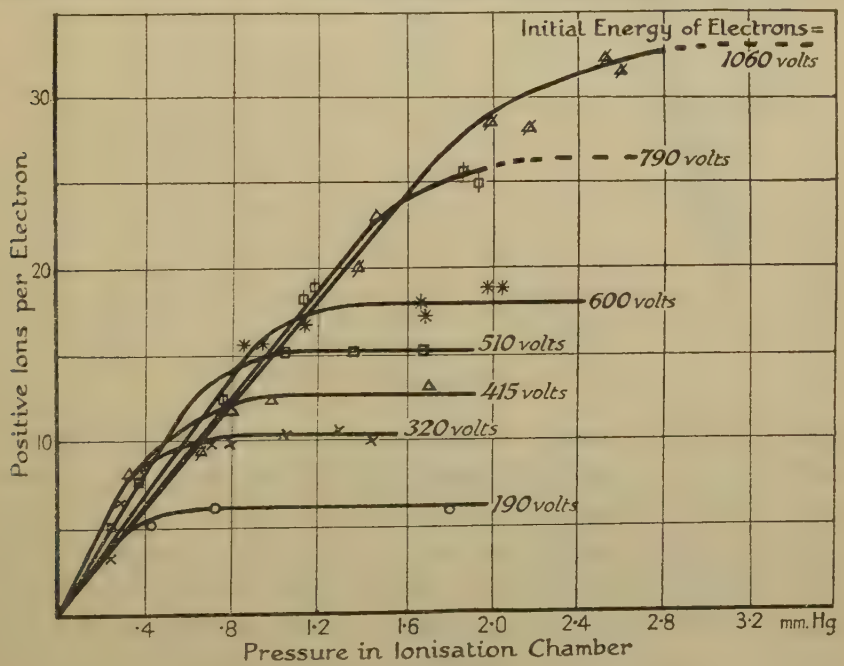


Fig. 1.—Ionisation in Helium.

the ionisation chamber. A typical set of measurements is given in Table I.; these are plotted in the 510-volt curve of fig. 1.

At low pressures the ionisation per electron was proportional to the pressures. At higher pressures it was constant and independent of pressure for any given initial energy. This constant ionisation measured the average number of ions produced by the complete absorption of an electron having the specified initial

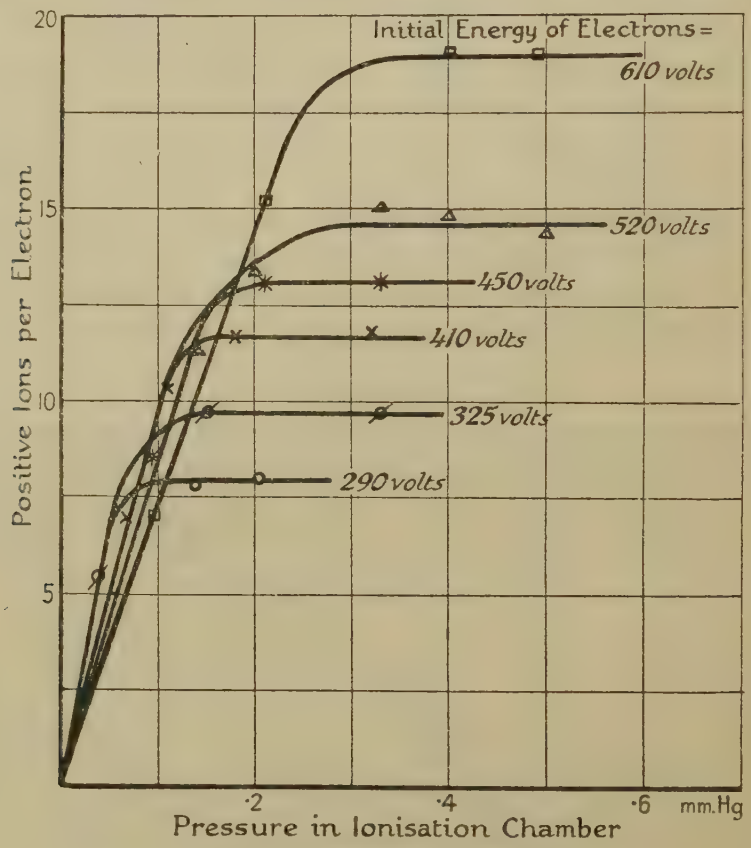


Fig. 2.—Ionisation in Argon.

energy. In fig. 5 the average number of positive ions formed per electron, for complete absorption, is plotted against the initial energy of the electron.

The pressures required to absorb the electrons in helium and in hydrogen were much greater than in other gases. Using a high-speed Gaede diffusion pump and large tubing between the filament chamber and pump, the highest pressure which could be maintained in the ionisation chamber was about 2 mm. Hg. This was inadequate to completely absorb the faster electrons in helium

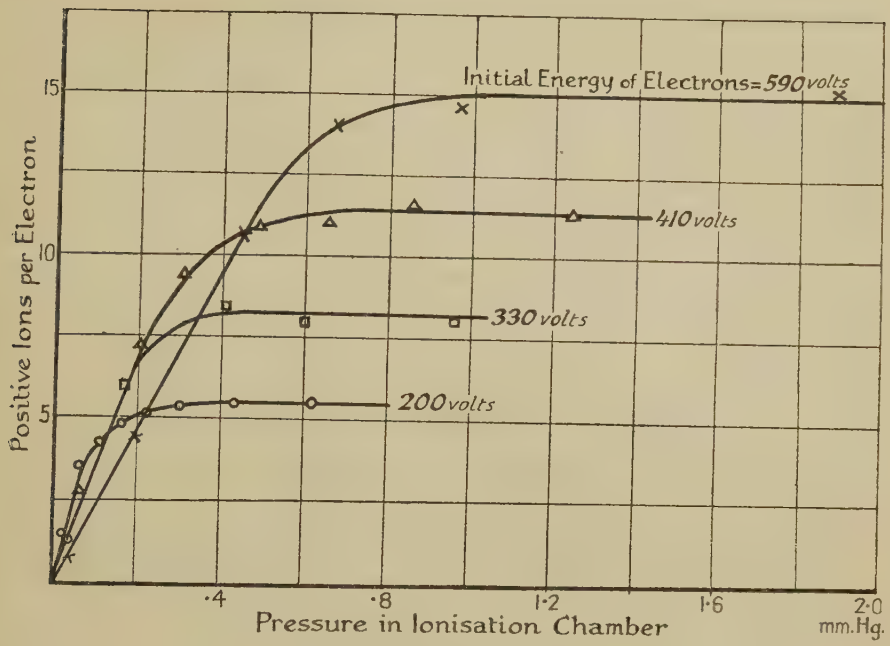


FIG. 3.—Ionisation in Hydrogen.

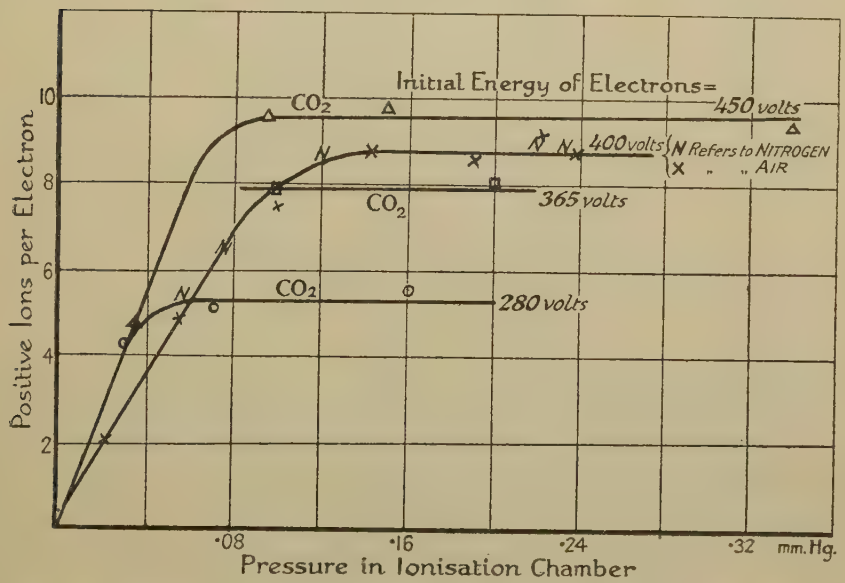


FIG. 4.—Ionisation in Carbon Dioxide and in Nitrogen.

or hydrogen. For this reason the curves of fig. 1, relating to the faster electrons in helium, are incomplete. The measurements in hydrogen were confined to

initial energies below 600 volts, for which adequate absorption could be obtained. In view of the results for helium and air there is little doubt that the line of fig. 5, referring to hydrogen, may be extrapolated at least to 1,000 volts.

Only one run was made in nitrogen, using an electron beam of initial energy equivalent to 400 volts. The results were identical, within the limits of measurement, with the results obtained in air under similar conditions. Further work on nitrogen was discontinued, as it was evident that any discrepancy between air and nitrogen was less than the experimental error. This result suggests that there can be no great difference between the ionisation in nitrogen and that in oxygen.

Table I.—Ionisation in Helium.

Initial velocity of electrons equivalent to 510 volts.

Electric current.	Positive ion current.	Positive ions per electron.	Pressure.
10^{-9} amps.	10^{-9} amps.		mm. Hg.
0·57	1·2	2·3	0·11
0·49	1·2		
0·56	2·8	5·0	0·24
0·57	2·8		
0·63	4·6	7·8	0·36
0·62	4·9		
0·55	7·6	13·7	0·78
0·56	8·0		
0·50	7·6	15·2x	1·10
0·55	7·9		
0·49	7·4		
0·47	7·6		
	Filament heating current reduced.		
0·34	5·3	15·2x	1·35
0·37	5·5		
0·30	4·6	15·3x	1·67
0·31	5·0		

Number of positive ions for complete absorption (mean of x) = 15·2.

Discussion.

1. *Energy Expenditure.*

The average energy expended in the formation of a pair of ions may be directly determined from fig. 5. This average energy, v, expressed in electron-volts is

$$v = \frac{\text{Initial energy of the electron stream (volts)}}{\text{Number of ion pairs for complete absorption}}.$$

Fig. 5 shows that v is sensibly constant, for any one gas, over the energy range examined, but varies in the different gases tested.

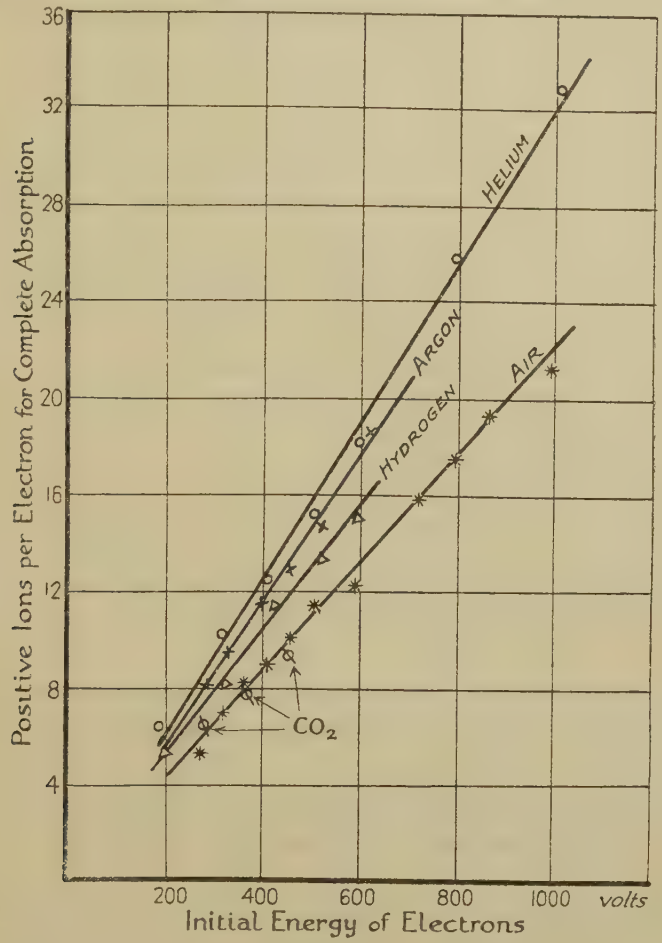


FIG. 5.

The ionisation potential p of a gas has been measured by determining the minimum energy with which a beam of electrons may form even an occasional pair of ions. This potential represents, in electron-volts, the energy associated with the separation of an electron from the normal atom without the transfer of an appreciable amount of kinetic energy to either of the ions so formed. An excess of the average energy v over the ionisation potential p must represent an expenditure of electronic energy by processes other than this mere division of a normal atom into positive and negative ions. The efficiency of ionisation F may be defined

$$F = p/v.$$

In the following table column 1 gives the types of ions which have been detected in the various gases; column 2, the values of p for the different ions; column 3, the values of v determined from fig. 5; and column 4 the values of F estimated from the lowest ionisation potential of the gas.

Table II.

Type of ion.	Ionisation potential.	Observed energy per pair ions.	Efficiency of ionisation.
		volts.	
He^+	24·5	31	0·78±0·05
*A^+	15·2	33	0·46
A^{++}	45·3		
H_2^+	15·9	37	0·43
N_2^+	17·0		
N^+	27·7	45	0·38
N^{++}	24·1		
CO^2	14·3	45	0·32

Note.—Ionisation potentials are taken from Bulletin of the National Research Council, U.S.A., No. 48, entitled 'Critical Potentials,' by Compton and Mohler.

* Henry A. Barton, 'Phys. Rev.,' Series 2, vol. 25, p. 469 (1925).

The efficiency of ionisation was, in all cases, much less than unity and varied from gas to gas. It is evident that ionisation was not the only process by means of which the initial energy of the electrons was dissipated. Further, since the efficiency of ionisation was lower in diatomic gases than in monatomic ones, it seems probable that some of the processes were associated with molecular rather than atomic structure.

The efficiencies given in Table II are calculated from the lowest ionisation potentials of the gas concerned. In many cases higher ionisation potentials exist, associated with the formation of doubly charged ions, ionisation of the K electron level, etc. To allow for these in the estimate of the efficiency of ionisation would require information concerning the relative probabilities of the different types of ions. Lacking precise information on this point, we can merely assume that the simplest type of ionisation is by far the most probable. As regards ionisation of the K level this assumption seems to be justified, for there is no appreciable change in the slope of the lines of fig. 5 as the initial energy of the electrons passes the critical potential associated with the K level. (The K level of oxygen and nitrogen and an L level of argon lie within the energy range investigated.) Some information concerning the relative proba-

bilities of the types of ions listed in Table II may be deduced from the work of H. D. Smyth and his collaborators. The significance of the critical potentials associated with these ions will be considered in the following discussion of the individual gases.

Helium.—The results obtained in helium are of especial interest as they refer to a monatomic gas of simple structure. There can be no dissipation of energy by molecular dissociation, and Smyth has found but a single ionisation potential at which He^+ ions are formed. Helium therefore appears to be the one gas to which the calculations of Mr. R. H. Fowler,* concerning ionisation by electronic impact in a gas having a single ionisation potential, are applicable.

Mr. Fowler assumed that at an ionising collision the primary electron loses energy equivalent to the ionisation potential of the ionised atom plus the kinetic energy with which the atomic electron is ejected from the atom. He estimated the velocity distribution of the secondary electrons by classical considerations. After an ionising collision both primary and secondary electrons may cause further ionisation provided their kinetic energy is greater than that equivalent to the ionisation potential. Residual kinetic energy less than this would be ineffective in ionisation. Mr. Fowler estimated that this residual energy would be about one quarter of the initial energy of the primary electrons, so that the efficiency of ionisation would be three-quarters.

This efficiency of ionisation is found experimentally to be $0{\cdot}78 \pm 0{\cdot}05$, a value in excellent agreement with Mr. Fowler's estimate. In the estimated efficiency no allowance was made for energy expended in excitation. The agreement between the estimated and experimental efficiency suggests that in helium there is but little energy expended in excitation.

Argon.—Argon represents a monatomic gas with a complex atomic structure. The efficiency of ionisation, as estimated from the lowest ionisation potential, is 0·43, much less than that found in helium, and slightly greater than that found in the diatomic gases tested. This estimate of the efficiency of ionisation should be corrected to allow for the doubly charged ions detected by Barton.† Barton's curves for 50-volt electrons indicated that the ratio of doubly charged to singly charged ions was about 10 : 1, which would make the efficiency of ionisation 0·46. It is possible that the relative number of doubly charged ions might increase with increasing energy of the ionising electrons, but even in the extreme case of all the ions being doubly charged the efficiency would be but 0·68, much lower than the efficiency in helium.

* R. H. Fowler, 'Proc. Camb. Phil. Soc.,' vol. 22, Pt. 5, p. 521 (1923).

† Barton, *loc. cit.*

It seems probable that in argon the initial energy of the ionising electrons may be partially dissipated by processes which are inappreciable in helium. A copious excitation of the argon atoms with the emission of a relatively intense radiation might constitute such a process, but a definite decision on this point must await further information concerning the relative probabilities of single ionisation, double ionisation, and excitation.

Hydrogen.—In hydrogen Smyth* has found ions with ratios of m/e corresponding to H_2^+, H^+, and H_3^+. Of these only H_2^+ was associated with a definite ionisation potential. The relative numbers of the three types of ions depended on the gas pressure, and Smyth concluded that the H^+ and H_3^+ ions were formed by dissociation of, and association with, the H_2^+ ions. It would thus appear that hydrogen is a gas with a single ionisation potential. In consequence we might expect an efficiency of ionisation of about 0·75 to correspond to Mr. Fowler's estimate of 3/4. The efficiency as determined experimentally is 0·43.

This marked discrepancy indicates that an appreciable fraction of the initial electronic energy is dissipated by processes other than ionisation. Excitation and dissociation are two processes which might absorb some electronic energy. The information available concerning these is scanty and rather conflicting. Smyth has found no evidence of dissociation accompanied by ionisation. Hughes,† on the other hand, has found a marked dissociation by electronic impact and, in view of Smyth's results, suggests that this may be accompanied by excitation instead of ionisation. His results indicate that the energy expended in dissociation with excitation may slightly exceed that spent in ionisation.

Nitrogen and Air.—Ionisation measurements in commercial nitrogen were found to agree with the corresponding measurements in air, to within the limits of experimental error. The results for air, published in a previous paper, are therefore taken as applicable to nitrogen.

In nitrogen Smyth‡ has found ions with ratio m/e corresponding to N_2^+, N^+, and N^{++}. The efficiency of ionisation in nitrogen, estimated from the ionisation potential corresponding to N_2^+, is 0·38 (Table II). This must be corrected to allow for the N^+ ions; the number of N^{++} ions detected by Smyth was too small to appreciably affect the efficiency. The relative number of N_2^+ and N^+ ions varied with the energy of the ionising electrons. Using 100-volt electrons, the ratio of molecular to atomic ions was 9 : 1. The corresponding efficiency of

* Smyth, 'Phys. Rev.,' vol. 25, p. 452 (1925).

† A. Ll. Hughes, 'Phil. Mag.,' vol. 48, p. 56 (1924).

‡ H. D. Smith, 'Roy. Soc. Proc.,' A, vol. 104, p. 121 (1923); and vol. 105, p. 116 (1924).

ionisation would be 0·40. For 600-volt electrons the ratio was 1 : 2, corresponding to an efficiency of 0·53. The correct efficiency is probably between 0·40 and 0·53.

Again, it seems probable that electronic energy was dissipated by processes other than ionisation. Hughes* has detected in nitrogen a profuse dissociation by electronic impact. It is possible that this process may absorb about as much energy as is absorbed in ionisation, but the data available does not permit a precise estimate.

Carbon Dioxide.—Smyth's method for determining m/e of the ions resulting from electronic impact has not, to the writer's knowledge, been applied to carbon dioxide. The only data available concerning this gas is, that it has an ionisation potential of 14·3 volts. The energy expended in the formation of a pair of ions was found to be 45 volts, indicating an efficiency of ionisation of 0·32. That this efficiency is lower than that found in any of the other gases tested is not surprising in view of the complexity of the carbon-dioxide molecule.

Summing up these speculations on the dissipation of the initial energy of the electrons we find—

In helium the experimental results are adequately accounted for by Mr. R. H. Fowler's views on ionisation by electronic impact in a gas with a single ionisation potential. The agreement between the experimental and estimated efficiency of ionisation indicates that but little energy is expended in exciting the helium atoms.

For the other gases tested the available information is insufficient to determine precisely the processes which contribute to the dissipation of the initial energy of the electrons. The results indicate that processes other than ionisation absorb appreciable amounts of the initial energy. It seems probable that in argon there is a copious excitation of the atoms. In hydrogen and nitrogen it is possible that dissociation, accompanied by excitation, may contribute largely to the dissipation of the initial electronic energy. The information concerning this process is rather conflicting, and a definite conclusion cannot yet be made concerning the extent to which dissociation contributes to the absorption of electronic energy.

* Hughes, *loc. cit.*

2. *A Comparison with Measurements on the Total Ionisation due to X-Rays and Alpha Rays.*

The ionisation produced by the complete absorption of X-rays in air has been measured by Kulenkampff.* The initial energy of the rays was measured by determining their heating effect when absorbed in lead. The accuracy to which the initial energy could be estimated was not specified. It was found that the average energy expended in forming a pair of ions was 35 volts. The discrepancy between this value and that of 45 volts found above for ionisation by electrons is rather striking. It is well beyond the limits of error of the present experiments, but may be associated with the determination of the initial energy of the X-rays.

The average energy expended in the formation of a pair of ions by an alpha particle has been measured in air by Geiger. Using Geiger's results for air, Gurney† has determined this energy expenditure in several gases. His results are given in the following table together with the corresponding energy expenditure by electrons, as deduced from figs. 1 to 5.

Table III.

	Ionisation potentials, P.	Average energy expended per ion pair.	
		α particle from Gurney, A.	Electron (from fig. 15), V.
	Volts.	Volts.	Volts.
Helium	24·5	26·2	31
Argon	15·2	24·0	33
Hydrogen	15·9	31·0	37
Air	17·0	33·0	45

From the data given in this table we may estimate the relative amounts of primary ionisation due to the alpha particle itself, and secondary ionisation due to the delta particles ejected by the alpha rays. At an ionising encounter the alpha particle must lose energy equivalent to the ionisation potential "p" plus the energy of the ejected delta particle. If ejected with sufficient energy, the delta particle may itself ionise with an average expenditure of v electron-volts per pair of ions. If the ratio of secondary to primary ionisation is n, on the average a delta particle will produce n ions involving an expenditure of nv electron-volts. This may be assumed to be the average energy with which the

* Kulenkampff, 'Ann. d. Physik,' vol. 80, p. 261 (1926).

† Gurney, 'Roy. Soc. Proc.,' A, vol. 107, p. 332 (1925).

delta particles are ejected. Thus, on the average, an alpha particle will expend $nv + p$ electron-volts per ionising encounter. As $n + 1$ positive ions result from this encounter, the average energy expended by the alpha ray per ion pair must be

$$A = \frac{nv + p}{n + 1}$$

or

$$n = \frac{A - p}{v - A}.$$

An estimate of n may be made for those gases in which the discrepancies between p, A and v are considerably greater than the experimental errors involved in the determination of these quantities.

In hydrogen $n = 2 \cdot 5$; in air $n = 1 \cdot 3$. These estimates are subject to errors of 30 or 40 per cent. They merely indicate that the secondary and primary ionisation of an alpha ray are of the same order of magnitude. In helium and argon the discrepancies between p, A and v are insufficient to permit even an approximate estimate of n.

3. *The Range of the Electrons.*

The curves of figs. 1 to 4 show two distinct stages. At low pressures the ionisation per electron is proportional to the pressure ; at higher pressures, adequate to completely absorb the electrons, it is constant. The transition between the two stages is fairly abrupt, and the critical pressure, just adequate to absorb the electrons, may be determined by extrapolating to their point of intersection the linear portions of the curves. Assuming that the lateral diffusion is small in comparison with the range of the electrons, the range at the critical pressure must be the length of the ionisation chamber, viz., 15 cm. If the range is inversely proportional to the pressure of the absorbing gas, at a standard pressure of 1 mm. Hg it would be

$$(15 \times \text{critical pressure}) \text{ cm.}$$

The ranges so estimated are plotted in fig. 6. The results for air are taken from a previous paper.

It is interesting to compare these ranges with the range as estimated from the rate of absorption of electronic energy. The rate at which the energy of relatively fast cathode rays is absorbed has been the subject of several investigations, both theoretical and experimental, but no measurements are available for the slow rays now under consideration. For these slow rays we have the

present measurements of the total ionisation and the measurements of the ionisation per unit path by Compton and Van Voorhis* and by Kossel.† The

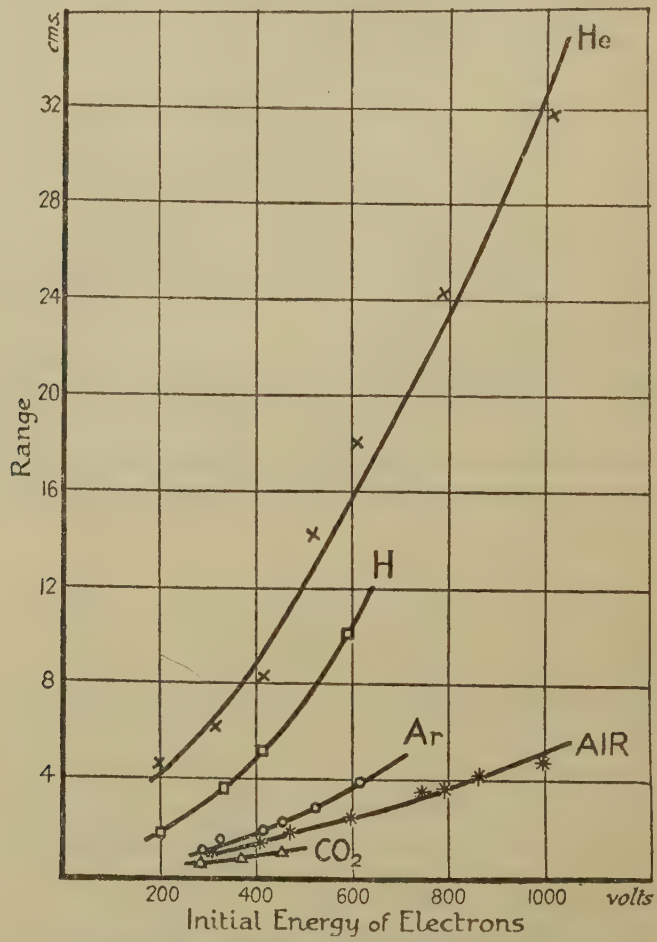


Fig. 6.—Range of Electron Beams (gas pressure = 1 mm. Hg).

rate of energy absorption necessary to correlate these measurements may be computed, and the range of the electrons estimated by integrating between energy limits equivalent to the initial energy of the electrons and the ionisation potential of the absorbing gas.

The calculations are confined to air, as it is only in that gas that both the total ionisation and the ionisation per unit path are known for the complete range of electron energies under consideration.

* Compton and Van Voorhis, 'Phys. Rev.,' vol. 27, p. 274 (1926).

† Kossel, 'Ann. d. Physik,' vol. 37, p. 393 (1912).

Let

I = the total ionisation due to the absorption of an electron which had an initial energy equivalent to V electron volts.

x = the distance the electron has travelled from an origin at which it had an energy of V electron volts.

q = the energy of the electron at a point x cm. from the origin.

i = the ionisation per unit path due to an electron with an energy q electron-volts travelling through air at a pressure of 1 mm. Hg.

p = the electronic energy equivalent to the ionisation potential of air.

R = the range of the electron, or that value of x at which q is less than p.

The number of primary ionising impacts in a distance dx is idx. At each impact a secondary electron would be ejected, and, if the average number of ions produced by each secondary is m, the ionisation in the distance dx is

$$i\,(l + m)\,.\,dx.$$

Measurements of the velocity distribution among secondary electrons ejected from metal foils shows that the great majority of these are ejected with energies inadequate to ionise air.* Thus the value of m is probably small in comparison with unity, and it has been neglected in the following computations. Integrating over the range of the electron we have

$$\mathrm{I} = \int_0^{\mathrm{R}} i dx$$

$$= \int_{\mathrm{V}}^{\mathrm{P}} i \frac{dx}{dq}\, dq.$$

Since no ionisation can occur when q is less than p,

$$\int_p^0 i \frac{dx}{dq}\, dq = 0,$$

therefore

$$\frac{d\mathrm{I}}{d\mathrm{V}} = -\, i \frac{dx}{dq}.$$

In air $d\mathrm{I}/d\mathrm{V} = 0{\cdot}022$ (fig. 5),

so that

$$\frac{dx}{dq} = -\,\frac{0{\cdot}022}{i} \text{ cm. per electron-volt.}$$

The experimental values of i are given in fig. 7 as a function q. The corre

* A. Becker, ' Ann. d. Physik,' vol. 78, p. 228 (1925).

sponding values of $-dx/dq$ are plotted in fig. 8. The range of the electrons for an initial energy V is

$$R = \int_{v}^{p} \frac{dx}{dq} dq = \int_{p}^{V} \frac{0{\cdot}022}{i} dq,$$

and may be obtained by a graphical integration of the curve of fig. 8. The values of the range are plotted in fig. 9. A comparison of fig. 9 with the air curve of fig. 6 shows a good agreement between the two estimates of the range of the electrons, and suggests that the measurements of total ionisation and of ionisation per unit path are consistent.

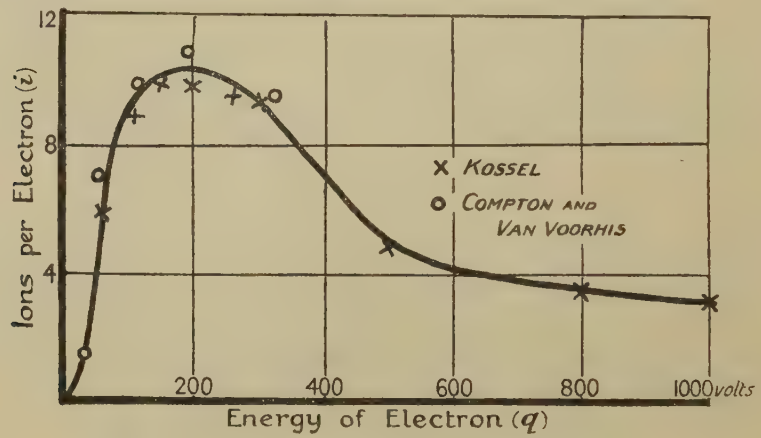


FIG. 7.—Ionisation per Unit Path in Air at 1 mm. Hg.

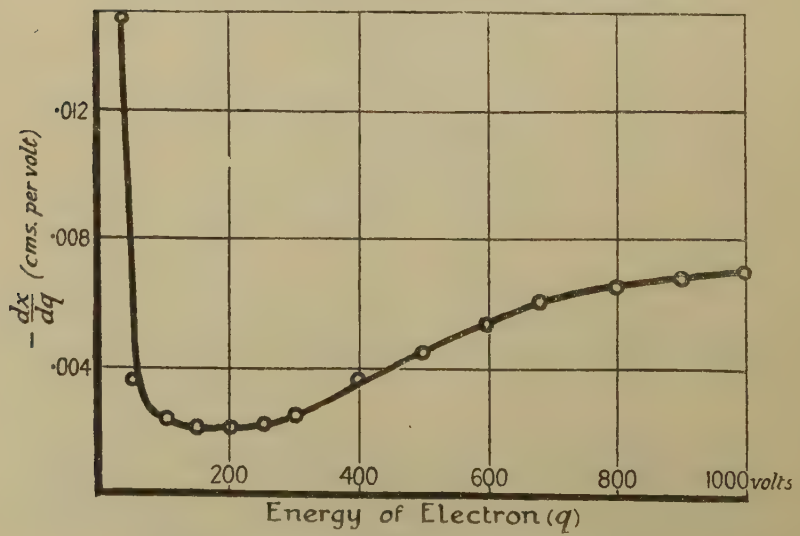


FIG. 8. $-\frac{dx}{dq} = \frac{0{\cdot}022}{i}$ in Air at 1 mm. Hg.

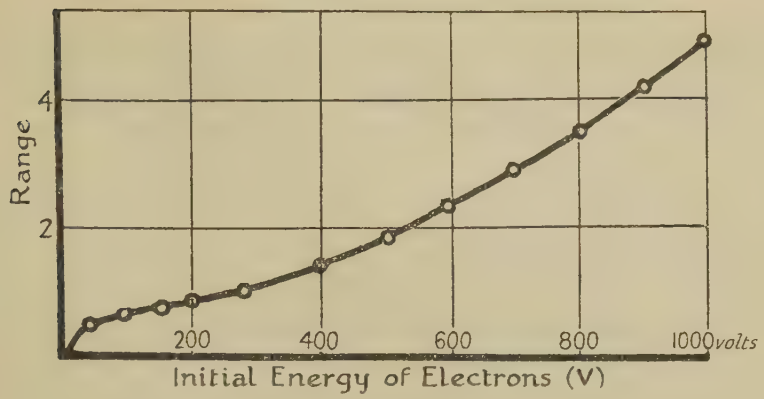


Fig. 9.—Range of Electrons in Air at 1 mm. Hg.

Summary.

The total ionisation due to the absorption of homogeneous beams of slow cathode rays has been measured in helium, argon, hydrogen, nitrogen, and carbon dioxide.

The average ionisation per electron for complete absorption was proportional to the initial energy of the cathode ray.

The average energy expended in the formation of a pair of ions and the efficiency of ionisation were as follows:—

	Energy per ion pair.	Efficiency.
Helium	31 electron-volts	0·78
Argon	33 ,,	0·46
Hydrogen	37 ,,	0·43
Nitrogen	45 ,,	0·38
Carbon dioxide	45 ,,	0·32

The rate of absorption of electronic energy necessary to correlate the present measurements of total ionisation with the measurements of Compton and Van Voorhis and of Kossel on ionisation per unit path has been estimated.

The range of the cathode rays, in air, was estimated

1. From the pressure necessary to absorb the rays within the length of the ionisation chamber.
2. From the rate of absorption of electronic energy.

The two estimates are concordant.

These experiments have been carried out under the supervision of Sir Ernest Rutherford. It is a pleasure to acknowledge my indebtedness to him and to express my thanks for his continued interest.

A Simple Radioactive Method for the Photographic Measurement of the Integrated Intensity of X-Ray Spectra.

By W. T. Astbury.

(Communicated by Sir William Bragg, F.R.S.—Received May 25, 1927.)

In this paper will be given a preliminary account of a method by which the intensity distribution in X-ray crystal photographs may be estimated by means of a simple and inexpensive apparatus. In the usual photometric procedure the blackening of a photographic plate is estimated by finding the intensity of light transmitted through it. In the method to be described the photographic plate is replaced by a carbon print, and the light and light-sensitive cell by α-rays and a simple α-ray electroscope respectively. The obvious advantages of such a system are, of course, ease of construction and negligible cost; but, as will be seen below, there already appears promise of a much more weighty advantage, namely, that with suitable adjustment the apparatus may be made to *integrate* the X-ray intensity of a crystal reflection after the manner of the Bragg ionisation-spectrometer.

Details are given below, but briefly the method is as follows. A carbon print is first made of the photographic *negative*. A carbon paper (or tissue, as it is called) consists merely of a pigmented film of gelatine sensitised with dilute aqueous potassium dichromate solution. On exposure to light such a film becomes insoluble, the depth of the insoluble layer after a given time being a measure of the intensity of the incident light. On development with warm water the remaining soluble portions are washed away, and there is left a film of thickness which varies according to the variation in blackness of the original negative. This insoluble film being very tough is readily detached, and, because it is so thin, its stopping-power for α-rays is just about of the right order. The variations in its thickness are thus very easily followed by aid of a few square millimetres of foil on which polonium has been deposited, a single slit, and the α-ray electroscope described by C. T. R. Wilson, which is usually figured in text-books. The latter can be constructed from an ordinary cigarette tin, and a simple tele-microscope with eyepiece scale.

In what follows, the apparatus and the results so far obtained with it will be described first. Finally will be given the details of the preparation of the "carbon" films.

Apparatus.

The plates used for the X-ray photographs were Ilford X-ray quarter plates. They were developed for 4 minutes at room temperature with the single-solution metol-hydroquinone developer (diluted to half strength) recommended by the makers on their plate boxes. The X-ray tube was of the Shearer type with copper anticathode.

The electroscope shown in fig. 1 was constructed of "tin" and to suit the short focus of the microscope available at the time, had the dimensions $7 \times 7 \times 2$ cm. The leaf was about $10 \times \frac{1}{2}$ mm. The bottom of the electroscope was closed with aluminium leaf equivalent in stopping-power to about 1 mm. of air.

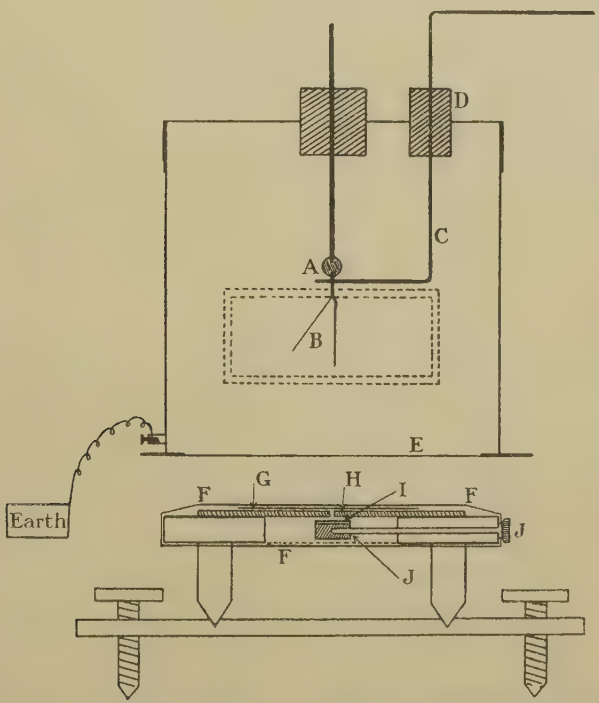


FIG. 1.—A, sulphur bead; B, leaf system; C, wire for charging leaf from rubbed ebonite or sealing wax; D, ebonite or rubber cork; E, aluminium leaf; F, elastic bands for holding carbon tissue in position; G, carbon tissue; H, glass slit system; I, deposit of polonium on metal foil; J, slider for polonium.

The slit system was of the simplest. It consisted merely of two pieces of glass (microscope slide) adjusted so that two edges were parallel and about $\frac{1}{2}$ mm. apart—actually the slit width was 0·43 mm. Its length was defined by two pieces of thick gummed paper, which was more than sufficient to stop the

α-rays. The polonium, deposited on a few square millimetres of silver foil, was mounted on a slider so that it could be withdrawn from under the glass slit when necessary ; it could always be replaced in the same position. The object of this arrangement was to enable the spectral lines of the X-ray photographs to be adjusted over the slit. Such lines are, of course, lighter than the background in the carbon print. The complete slit system could be removed from under the electroscope and held up to the light. By this means the film was adjusted simply by hand and afterwards held in position by two thin elastic bands. When the required line had been brought over the slit, the whole could then be replaced under the electroscope, and always in the same position. After this procedure had been carried out, the measurement consisted only in observing the time of fall of the leaf over 100 divisions of the eyepiece scale of the microscope. (The mean of several readings must always be taken, on account of the probability variations in the α-ray emission.) The reciprocal of this number (the natural leak of the electroscope was negligible) was taken as representing the intensity of α-ray ionisation for the thickness of carbon tissue under examination. Thus to every value of the X-ray intensity on the original negative there corresponded a certain α-ray intensity.

Preliminary considerations.

The maximum thickness of the insoluble gelatine film in a carbon print is of the order of 1/50 mm., equivalent in stopping power to about 1·5 to 2 cm. of air when tested with the electroscope. The film is practically inextensible and, considering its thinness, very tough—properties which allow it to be easily stripped from its support without injury. It can be seen at once that, after development, the thickness of a film for a given light intensity during printing should be independent of irregularities of thickness in the original undeveloped film, provided that the pigmentation and sensitising are uniform. This is a remarkable property which has naturally contributed a great deal to the success of the measurements being described. It arises from the fact that, before development with warm water, the film is transferred to a temporary support, so that the still soluble gelatine may be washed away from the side of the film *opposite* to that which has been exposed to the printing light. Thus for a given intensity of printing light a constant thickness of insoluble gelatine remains behind after the washing has been completed.

This point was tested by exposing a number of films under plain glass. Either sunlight or arc-light may be used, though the former on bright days is much quicker. The exposure is controlled by an " actinometer "—a box in which

a small area of printing-out-paper can be exposed simultaneously, the unit of exposure being that required to darken the p.o.p. to a standard tint. The number of units most suitable for a given negative is soon found by trial. It was found that with photogravure tissues (G 12), developed as described below, the electroscope readings were constant over the central area of the film (about 20 sq. cm. were examined) to within about 1 per cent., in fact the variations were of the order of the probability variations of the α-ray emission. This was a very encouraging verification of the prediction mentioned above and showed that quite high accuracy might be expected by taking the mean from several prints.

The next step involved an X-ray photograph the intensity variations on which were known. For X-rays the photographic effect of twice the intensity can be obtained very closely by doubling the time of exposure,* but it was found more convenient to adopt the following procedure. The X-rays passed through a slit about 2 cm. long and fell upon a selected strip of mica of similar dimensions. The CuK_α and K_β lines produced on a photographic plate 10 cm. distant were then several centimetres long and nearly 1 mm. broad. It was then assumed that the central portion (1·5 cm., say) of these lines was practically uniform in intensity. On examining carbon prints of such a photograph with the electroscope, the glass slit being 3 mm. $\times$ 0·43 mm., it was found that this was indeed the case, since the variations in electroscope readings were of the same order as had been observed in the plain films mentioned above.

This result afforded an easy method of calibrating an X-ray photograph. As is well known, the only sound method of using photographic plates for accurate intensity measurements involves a calibration wedge the history of which is identical with that of the plate under examination. In other words, whatever photometric system is used, the basic step in the measurement must always be the finding of two points on the *same* plate where the photographic effects are identical. This means that *every photographic plate must bear its own calibration wedge which has passed through exactly the same stages of development, fixing, etc., as have been suffered by the photographic lines or spots to be measured by aid of the wedge.*

In these experiments the calibration wedge was made by adding successive thicknesses (1/1000 in.) of aluminium foil. Such a step-wedge was constructed with steps 2 or 3 mm. long and fixed to the paper covering the X-ray plate over the area where the third order CuK_α mica line was known to be produced. The wedge thus overlay the central portion of the line which had previously

* A. Bouwers, 'Z. f. Physik,' vol. 14, p. 174 (1923); 'Physica,' vol. 3, p. 113 (1923); and vol. 5, p. 8 (1925); 'Dissertation Utrecht,' 1924.

been shown to be approximately uniform in intensity. The values of X-ray intensity corresponding to the steps of the wedge were then calculated from the thicknesses of the steps, the angle of incidence of the rays on the plate and the absorption coefficient of CuK_a rays in aluminium. The wedge so obtained served for the measurement of other lines on the same plate, since its photographic history throughout was identical with theirs.

Experimental Results.

The first quantitative measurements were made with the mica photographs mentioned above. These showed the first three orders of reflection from the cleavage plane. A portion 1·2 cm. long of the third order K_a line was used for calibration, this being divided into four steps of 3 mm. each, corresponding to the X-ray intensities, 100, 71·4, 51·0 and 36·4. In these first experiments the slit system was placed at an arbitrary distance (generally 1–2 cm.) below the electroscope, but, as will be seen below, this is not the best method of using the apparatus. The width of the slit was about half the width of the line, and it was assumed that when the maximum of the line covered the slit, the apparatus was measuring a small area (3 mm. × 0·43 mm.) of approximately uniform intensity. The α-ray intensities of the four steps of the wedge were then plotted against the calculated X-ray intensities corresponding to them. The four points so obtained lay approximately on a straight line. Readings for the central 3 mm. of the third order β-line and the first and second order α-lines were then taken, and also in each case the readings for the photographic background on both sides of the line. The corresponding X-ray intensities were then read off from the calibration curve. The results for five different prints of the first photograph were as follows :—

Table I.

3rd order.	Background.	3rd β.	Background.	2nd order.	Background.	1st order.	Background.
100	41	53	41½	54	35	78	34
100	41½	48½	39½	57	37½	77½	37½
100	42½	47½	39½	54	39	74½	37
100	42	50	39	54	36½	78	37
100	41½	47½	39	55½	38	74½	34½
100	41½	49½	39½	55	37	76½	36 Mean

The effects of the respective backgrounds having been subtracted, the corrected relative intensities were then :—

$$\frac{\text{3rd order}}{58\frac{1}{2}}, \quad \frac{\text{3rd order } \beta}{10}, \quad \frac{\text{2nd order}}{18}, \quad \frac{\text{1st order}}{40\frac{1}{2}},$$

which gave for the ratio K_α/K_β, $58\frac{1}{2}/10 = 5{\cdot}85$.

As a control upon this result another mica photograph was taken and measured up in a similar manner. The values obtained from three prints were as follows :—

Table II.

3rd order.	Background.	3rd β.	Background.	2nd order.	Background.	1st order.	Background.
100	40	44	33	48	$28\frac{1}{2}$	$72\frac{1}{2}$	32
100	$39\frac{1}{2}$	$44\frac{1}{2}$	34	$50\frac{1}{2}$	$31\frac{1}{2}$	$69\frac{1}{2}$	$33\frac{1}{2}$
100	41	$46\frac{1}{2}$	$35\frac{1}{2}$	$46\frac{1}{2}$	29	67	$32\frac{1}{2}$
100	40	45	34	$48\frac{1}{2}$	$29\frac{1}{2}$	$69\frac{1}{2}$	$32\frac{1}{2}$ Mean

Hence

$$\frac{\text{3rd order}}{60}, \quad \frac{\text{3rd order } \beta}{11}, \quad \frac{\text{2nd order}}{19}, \quad \frac{\text{1st order}}{37},$$

and the ratio K_α/K_β, $60/11 = 5{\cdot}45$.

This agreement was exceedingly satisfactory, for, as can be seen from the results tabulated above, it was obtained under relatively bad conditions, since the general X-ray background was comparable with the reflections themselves. The means of the two sets of observations give for muscovite mica, 1st order : 2nd order : 3rd order = 66 : 31 : 100, and for copper, $K_\alpha/K_\beta = 5{\cdot}65$. This last value must be corrected for the differential absorption of K_α and K_β in the window of the X-ray tube. The final corrected value for $CuK_\alpha/K_\beta = 6{\cdot}2$. (The window was aluminium foil, 1/1000 in. thick.) This result lies between the two values quoted by Siegbahn,* viz., 6·53, found by Unnewehr, and 6·0, found by Siegbahn and Zácek. It is based on the assumption, which must be nearly true, that K_α and K_β are equivalent in their effects on the photographic plate. Of course, the intensities measured above are not the true integrated intensities required for accurate X-ray analysis. But the edges of the lines were so sharp that the ratios obtained must be very nearly correct. A source of error is provided by the nature of the general X-ray background, always an annoying feature in the accurate intensity measurement of crystal reflections.

* M. Siegbahn, "Spectroscopy of X-rays," English translation, pp. 96–97.

In the method used here it is assumed that the background consists mostly of copper radiation ; otherwise the calibration by an aluminium absorption wedge would be invalid. The very satisfactory value obtained for the ratio K_α/K_β indicates that the assumptions involved in its measurement are not seriously wrong. But the only safe way of eliminating the uncertainty of background values is to reduce them to a minimum. In the observations to be described next, this precept was put into practice.

The experiments outlined above must be regarded as rather of a tentative nature. A much stricter test was necessary, over a wider range of intensities. There are always two main difficulties : 1, the background, and 2, the difficulty of obtaining control measurements by other methods. The first was practically eliminated by using a very small oscillation (5°) instead of the 20° used above. This reduced the background to less than 10 per cent. of the line, and the reflection of other wave-lengths on to the area occupied by the line must have been very small. The second obstacle is bound up with the nature of crystal reflection itself, the intensity of which has been shown to depend upon a number of factors which are often extremely difficult to evaluate. On this account it was considered advisable to use some method which would be independent of the theory of crystal reflection. For this purpose the following procedure was adopted.

The same CuK_α mica third order reflection was used as before, with an aluminium step-wedge of seven steps each 2 mm. long, corresponding to the X-ray intensities, 100, 71·4, 51·0, 36·4, 26·0, 18·5, and 13·2. Two lines, covered in turn by this step-wedge, were made upon the photographic plate in the two positions for which the reflected X-rays are equally inclined to the normal to the plate. The time of exposure for one line was 25 minutes and for the other 20 minutes, the X-ray tube running very steadily all the while. Thus were produced on the plate two lines that were each divided into seven gradations of intensity, and that were exactly alike in every way except in that one was only 4/5 as intense as the other. The plan was to use one wedge as a calibration wedge and to find from it the relative intensities in the other. The intensity ratios in the two wedges should be the same, of course, and furthermore we should find the ratio 4 : 5 between the intensities of corresponding steps. Table III shows the results from seven prints of this photograph. The first four are photogravure tissues G 12 and the last three G 2 (*see below*). The X-ray intensity of the strongest step of the more intense wedge is counted as 100.

Table III.

1st step.	2nd step.	3rd step.	4th step.	5th step.	6th step.	7th step.
81	55	$38\frac{1}{2}$	$28\frac{1}{2}$	$20\frac{1}{2}$	$14\frac{1}{2}$	$11\frac{1}{2}$
$78\frac{1}{2}$	$55\frac{1}{2}$	$39\frac{1}{2}$	29	$21\frac{1}{2}$	15	$11\frac{1}{2}$
$81\frac{1}{2}$	$53\frac{1}{2}$	$38\frac{1}{2}$	27	21	16	$12\frac{1}{2}$
80	55	$38\frac{1}{2}$	$28\frac{1}{2}$	$21\frac{1}{2}$	$15\frac{1}{2}$	$12\frac{1}{2}$
$80\frac{1}{2}$	$57\frac{1}{2}$	$40\frac{1}{2}$	29	21	$14\frac{1}{2}$	$11\frac{1}{2}$
79	55	$37\frac{1}{2}$	$27\frac{1}{2}$	21	14	11
80	$55\frac{1}{2}$	$38\frac{1}{2}$	*	$19\frac{1}{2}$	14	10
80	$55\frac{1}{2}$	39	$28\frac{1}{2}$	21	15	$11\frac{1}{2}$ Mean

* Film injured here.

The intensities of the steps of the weaker wedge are therefore in the ratio, $80 : 55\frac{1}{2} : 39 : 28\frac{1}{2} : 21 : 15 : 11\frac{1}{2}$, *i.e.*, $100 : 69\frac{1}{2} : 49 : 35\frac{1}{2} : 26\frac{1}{2} : 19 : 14\frac{1}{2}$. A comparison of these figures with those calculated from the constants of the wedge

Calculated $100 : 71{\cdot}4 : 51{\cdot}0 : 36{\cdot}4 : 26{\cdot}0 : 18{\cdot}5 : 13{\cdot}2$

Observed $100 : 69\frac{1}{2} : 49 : 35\frac{1}{2} : 26\frac{1}{2} : 19 : 14\frac{1}{2}$,

must be considered very satisfactory, if we bear in mind the comparative crudeness of the apparatus and technique described above. In any case, perfect agreement is to be expected only if the X-ray negative is uniform and uniformly developed, the calibration line is uniform in intensity and the aluminium foil of the wedge is of uniform thickness.

It is also satisfactory to note that the ratio 4 : 5 is found for the intensities of corresponding steps of the two wedges. The intensity of the 7th step was obtained by extrapolation, and perhaps on this account is not so accurate as the intensities of the other steps (see below). A typical pair of curves (the 5th in Table III) for the two wedges is shown in fig. 2.

The lower one is the calibration curve obtained by measurement of the more intense of the two photographic wedges. The ordinates of the weaker wedge cut this calibration curve in the values of X-ray intensity, $80\frac{1}{2}$, $57\frac{1}{2}$, $40\frac{1}{2}$, 29, 21, $14\frac{1}{2}$, and $11\frac{1}{2}$. The abscissæ are the reciprocals of the times of fall of the leaf over 100 divisions of the eyepiece scale, multiplied by 1000, while the ordinates are in arbitrary units.

The above measurements were obtained by a procedure which differed somewhat from that used in the case of the mica photographs. In the latter the slit system was placed at an arbitrary distance below the electroscope, but for the double wedge investigation the α-ray intensity corresponding to zero X-ray intensity was made zero also. This was accomplished by lowering the slit

system until the electroscope leaf showed no deflection for the area of the print corresponding to those parts of the original negative which had been covered

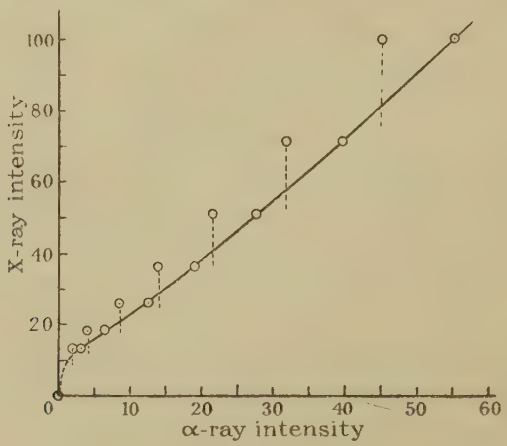


Fig. 2.

by lead. It is clear that measurements obtained in this fashion are more accurate than those obtained with the slit near to the bottom of the electroscope, because the α-rays have short and definite range (about 3·8 cm. for polonium). The percentage increase in ionisation for a given decrease in thickness of the carbon film becomes smaller and smaller as the slit system is brought nearer to the bottom of the electroscope.

Temperature Effect.

In these first experiments with the radioactivity photometer the effect of temperature was neglected. The only precaution taken was to carry out the observations in a room which was free from draughts and in which the temperature conditions had become fairly steady. But as a matter of fact temperature changes may have an important effect on the readings of the weak intensities, because the range in air of the α-particles increases from 3·721 cms. at 0° C. to 3·925 cms. at 15° C. Thus a change in temperature alters that distance of the slit system below the electroscope which corresponds to zero α-ray intensity, and therefore alters the proportion of the total range which, for a given X-ray intensity, is included inside the electroscope.

This trouble may be easily eliminated by arranging that the α-rays, before they enter the electroscope, do not pass through air but through a medium—*e.g.*, metal foil—of which the coefficient of expansion is small compared with that of a gas. Then, since the electroscope is more than deep enough to absorb

the rest of the range, the apparatus becomes practically independent of ordinary temperature changes. For this purpose a slight modification of the apparatus sketched in fig. 1 is necessary. As the maximum stopping power of the carbon tissue is about that of 2 cm. of air, the bottom of the electroscope may be closed permanently, not with aluminium leaf, but with foil of, say, 1 cm. stopping power. Afterwards, successive thicknesses of leaf may be temporarily attached to this instead of lowering the slit system. Zero deflection for zero X-ray intensity can then always be obtained with the slit system very close to the bottom of the electroscope.

The Integrating Property.

As mentioned above, the property of α-rays of having a short and definite range gives us at once the power of making, by a small adjustment of the apparatus, the X-ray and α-ray intensities vanish simultaneously. This means that the radioactivity photometer under discussion possesses in itself the first indispensable characteristic of an integrating photometer. The second condition that must be satisfied is that the α-ray intensity must always be directly proportional to the X-ray intensity. Fig. 2 shows how nearly this condition is already satisfied in the method of measurement adopted for the double wedge investigation. The curves shown are not far from straight lines through the origin. But as they stand, they have one unsatisfactory feature that must be remedied, and that is the slope of the part near the origin. Both curves are based on the adjustment of zero α-ray intensity for zero X-ray intensity, and so it is clear that there must be a sharp change in curvature quite close to the origin. Such a discontinuity is fatal for the integrating property. Evidently, it is a consequence of the shape of the Bragg α-ray ionisation curve.

It is possible to draw an interesting conclusion from a simple mathematical consideration of the question. Let I_0 be the intensity of the printing light and I the intensity that is transmitted through the negative at the point where the X-ray intensity is X. Then, since it is known that for small X-ray intensities the photographic density is proportional to the X-ray intensity, we have

$$X = A \log (I_0/I) \qquad (1)$$

where A is a constant. The light intensity I will fall on the carbon tissue and be absorbed, presumably, according to the ordinary absorption law. But there will be a minimum intensity i that is necessary to bring about the reaction between the potassium dichromate and gelatine which results in the insolu-

bility of the latter. Thus a thickness t is made insoluble, and, if μ is the absorption coefficient,

$$i = \mathrm{I}e^{-\mu t}. \qquad (2)$$

From (1) and (2) we have

$$t = 1/\mu\,(\log \mathrm{I}_0/i - \mathrm{X}/\mathrm{A}).$$

Hence, if T is the thickness of insoluble carbon tissue corresponding to zero X-ray intensity,

$$\mathrm{X} = \mu\mathrm{A}\,(\mathrm{T} - t). \qquad (3)$$

Or, in other words, the X-ray intensity is proportional to the decrease in thickness of the insoluble carbon tissue; that is, the X-ray intensity at any point of the carbon print is measured by the depth of the hollow there below the level corresponding to zero X-ray intensity. If this deduction is sound, it means that if we were to explore the contour of the print with some delicate device such as an optical lever, it would trace out peaks of X-ray reflection exactly of the same shape as those that may be obtained on a Bragg ionisation spectrometer.

In terms of electroscope readings this result may be expressed by saying that the X-ray intensity at any point of the print is proportional to the additional length of range that enters the electroscope when that point lies over the slit. Thus, for small X-ray intensities, the curve showing the increase of α-ray ionisation with increasing X-ray intensity should be simply the curve which shows the increase of α-ray ionisation as we decrease by successive equal amounts the thickness of film covering the radioactive source. Such a curve was obtained long ago by Rutherford*; it is approximately logarithmic, following near the origin a course similar to that of the curves of fig. 2.

Owing to the difficulty of obtaining thin uniform films of 1 or 2 mm. stopping power, an analogous experiment was made by decreasing 1 mm. at a time the distance between the slit and the electroscope. The increase in ionisation is then faster than in Rutherford's experiment owing to the increase in solid angle of the α-ray beam, but the result is sufficiently close for the purpose. It showed that equation (3) above is indeed approximately true for small photographic densities, up to about 0·4,† as far as could be judged. At higher densities the wedge and electroscope curves diverge, the latter being slightly concave downwards. This is quite what would be expected, for two reasons: 1, the increase of solid angle mentioned above, and 2, the photographic density increases more slowly with strong X-ray intensities (see Bouwers, *loc. cit.*) than is demanded

* 'Radioactive Substances and their Radiations,' p. 146.

† For measurements of photographic density I am indebted to Dr. G. Shearer, of the National Physical Laboratory.

by the straight line law which holds for weak intensities; that is, the carbon film is always thicker and therefore the ionisation smaller than the linear law requires.

The above discussion throws a considerable light on the question of the way the radioactivity photometer must be used in order that it may act also as an integrating photometer. It is clear that the curve obtained from the step-wedge is a compromise between the photographic blackening curve and the Bragg curve which shows the variation in ionisation along the path of the α-particles. The short steep slope near the origin in the curves of fig. 2, which was the cause of some uncertainty in the extrapolated value of the intensity of the seventh step of the weaker wedge, obviously corresponds to that part of the Bragg curve, known as the "tail," which is due to "straggling" of the α-particles. It can be eliminated in a very simple fashion by arithmetical subtraction of the ionisation due to it. In other words, what we have to do is to adjust the slit-electroscope distance so that zero X-ray intensity corresponds to the foot of the steep part of the Bragg curve, and then to subtract the α-ray intensity corresponding to zero X-ray intensity from all other readings. This procedure makes the curve pass through the origin and at the same time removes the discontinuity shown in fig. 2. The "tail" of the Bragg curve is somewhat indefinite, and so it is necessary to find the "optimum" position of the slit. This has been done for the weaker wedge, with the following results (Table 4). The five columns correspond to five positions of the slit, the first being the position where zero X-ray intensity produces zero α-ray intensity, and the other four positions being at 2, 3, 4 and 5 mm. nearer the electroscope respectively. The "tail effect" has been subtracted in each case (it is zero in the first, of course), and then the results have all been reduced so as to make the α-ray intensity of the third step always equal to 51.

Table IV.

Distance of slit from zero position :	0 mm.	2 mm.	3 mm.	4 mm.	5 mm.
1st step	103	92½	91	87½	86
2nd step	73½	68	67	66	64
3rd step	51	51	51	51	51
4th step	33½	36½	36½	37½	38½
5th step	21	25	26½	27½	28
6th step	10½	15½	16½	18	18
7th step	6	9½	10	12	12½

These results are only rough preliminary values, but they are sufficiently accurate to make the point here desired; and that is, that the radioactivity photometer with a certain adjustment may be made to function as an integrating photometer over quite a reasonable range. The calculated intensities of the seven steps of the wedge are 100, 71·4, 51·0, 36·4, 26·0, 18·5 and 13·2, and the numbers reproduced in the above table are readings of the seven corresponding α-ray intensities obtained without reference to the calibration curve. It will be seen how, near the zero position, the stronger intensities are in agreement, while raising the slit 4–5 mm. brings the smaller intensities into correspondence. The values 51, $37\frac{1}{2}$, $27\frac{1}{2}$, 18, 12 and 0 obtained at 4 mm. are in very gratifying agreement with the X-ray intensities 51, 36·4, 26·0, 18·5, 13·2 and 0. The curves for the zero position (I) and for 4 mm. (II) are shown for comparison in fig. 3.

The photographic density of the most intense step (X-ray intensity = 100)

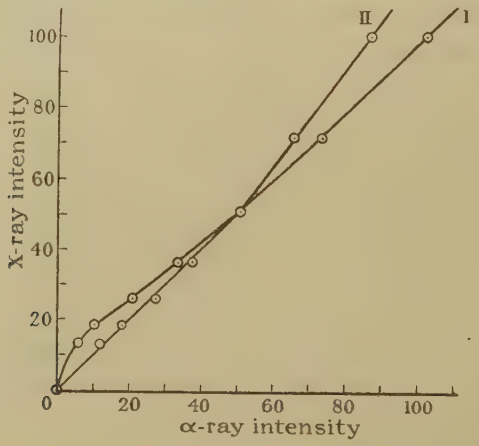


Fig. 3.

of the wedge under discussion was 1·33, and for the third step (X-ray intensity = 51) it was 0·81. Now 0·8 corresponds to quite a considerable blackening of the photographic plate, yet up to this density the X-ray intensities and α-ray intensities are closely proportional. Hence, even at this preliminary stage, it is apparent that the radioactivity photometer may be used as an integrating photometer for most of the spots on an X-ray photograph. And, after all, it is easy enough to cut down the time of exposure of an X-ray plate so that the photographic densities may fall within the prescribed limits.

But we have not yet exhausted all the means of improving the technique and of extending the integrating range still further. As pointed out above

the photometer curve is a compromise between the photographic density curve and the Bragg ionisation curve, and a study of Table IV indicates at once how to make the best of this circumstance. It shows that curve II turns up for higher intensities chiefly because the maximum of the ionisation curve has been reached. Fig. 4 shows schematically what is happening. In the diagram the slit-electroscope distance is supposed adjusted to (say) the 4 mm. position, so that the "tail" is just inside the electroscope. As the X-ray intensity increases, the carbon film becomes thinner and more and more of the Bragg curve enters the electroscope, the α-ray intensity observed being the area under the part of the curve that has entered. The α-ray beam is very divergent, so that the part between the tail and the maximum is flattened somewhat (apparently with the apparatus used here it corresponds to about 1·5 cms. of the range). Evidently then, what we have to do is to cut down the maximum thickness of the carbon print so that its stopping-power does not exceed the stopping-power of this convenient part of the α-range. And, of course, this can be done simply by decreasing the printing exposure.

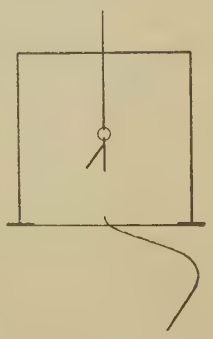


Fig. 4.

Another possible way of improving the method is by a more suitable choice of photographic plate. A plate whose density curve approximates more nearly to a linear law than the one used in these experiments, would probably give a more exact linear relation still for the radioactive photometer and over a wider range.

But these points remain to be examined in detail by further experiments. The question, too, of the polonium deposit has still to be gone into. For finding the integrated intensity of spots a uniform deposit will be necessary. This will then be placed very close under an aperture which is larger than the spots to be examined. If the active deposit is larger than the aperture and the polonium and carbon tissue are very close to each other, the ionisation effects due to the different parts of the spot will be approximately additive. The best position of the aperture under the electroscope having been determined for a given type of photographic plate and development, it should then be possible to keep this arrangement once and for all for plates of the same make. It is hoped to continue this investigation along these lines.

The intensity of reflection of X-rays from crystals is in general so sensitive to changes of interatomic distances that comparatively rough measurements

of intensity are very valuable for the determination of these distances. On this account especially it is the writer's hope that the cheap and easily constructed apparatus described above will prove of great use to workers whose researches would be aided by intensity measurements, yet who are not in a position to acquire one of the usual photometers, which are generally so expensive. Whatever the outcome of a further examination of the integrating property, the experimental results communicated above clearly demonstrate that the radioactivity photometer is capable of yielding very satisfactory intensity values when the usual photometric technique is employed.

*Preparation of the Films.**

The carbon tissues found so far to give the best results are the ordinary photogravure tissues† manufactured for the printing trade. The tissue is sensitised in the morning by immersion in a $1\frac{1}{2}$ per cent. aqueous solution of potassium dichromate (made just neutral with ammonia) for about 2 minutes, care being taken to see that all air bubbles are removed by means of a camel hair brush. The sensitising can be carried out in red, weak electric, or gas light. After removal from the sensitising bath the tissue is carefully squeegeed with a flat rubber squeegee on to a clean glass plate and placed in a dark cupboard with a fan. Drying should be completed in a few hours, in fact *prolonged drying should be avoided.* If needed for ordinary photographic work the sensitive films may then be kept in a light-tight box for perhaps a week, but for the purposes of photometric measurements they must be used as soon after drying as possible—hence the recommendation to sensitise them in the morning so that they may be developed in the afternoon. This precaution is necessary because after a day or so they commence to become insoluble of themselves, without the apparent intervention of light.

For printing, an actinometer is used, the negative being provided with a "safe-edge" of black paper so that a soluble border of about $\frac{1}{4}$ to $\frac{1}{2}$ inch is left on the print. It is usual to place a rubber pad behind the tissue while it is in the printing frame, in order to exclude damp.

For development, a convenient device has been adopted to ensure complete and uniform development and to permit easy stripping of the film when dry.

* In the preparation of these films I have been assisted throughout by Mr. H. Smith, of the Davy Faraday Laboratory.

† All the tissues and photographic equipment and materials used in these experiments are obtainable from the Autotype Company, 59, New Oxford Street, London. The same firm also publishes a small book, "A.B.C. Guide to Autotype Carbon Printing," in which will be found a complete description of the process.

This consists of a rectangle of metal as a back support and a frame of thin metal to act as a border to the print to hold it down on to the back support. It is used as follows. The metal border is first covered on one side with ordinary *single* transfer paper, using shellac as the adhesive. The development of the exposed tissue is then proceeded with by first placing it in cold water together with a piece of *temporary* transfer paper that has been previously waxed.* In a short time it will be noticed that the tissue begins to curl up ; *the moment it begins to uncurl*, it must be brought face down on to the temporary transfer paper, both removed from the water and squeegeed together, firmly but not roughly, by aid of a flat rubber squeegee and (say) a piece of plate glass. This whole operation should take some couple of minutes. After squeegeeing, the two adhering papers are placed between blotting paper under a flat weight for 20 minutes. They are then placed in warm water at about 40° C., when in a short time the soluble gelatine will begin to ooze out. When it appears, the paper backing of the tissue can be stripped off, leaving the film adhering to the waxed transfer paper. This is then placed on the rectangular metal support, covered with the metal border and clipped at the corners, after which development with the warm water can be proceeded with without fear of the film becoming detached. The appearance of this developing frame is shown in fig. 5.

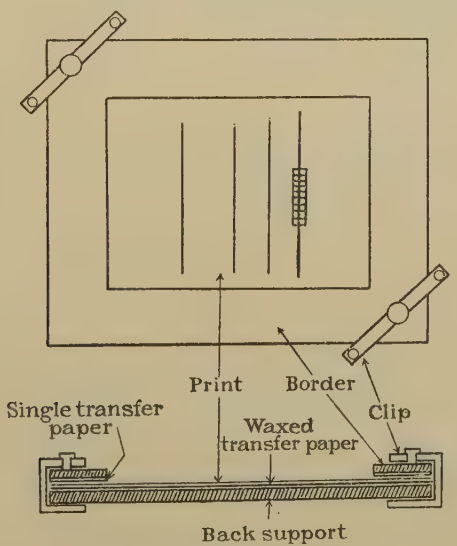


Fig. 5.

* The waxing solution consists of about 3 gm. resin and 1 gm. beeswax dissolved in some 100 c.c. of turpentine. See literature supplied by Autotype Company for method of waxing transfer papers. The minimum quantity of solution should be used.

It is best to have a continuous stream of warm water running into the developing dish. For accurate intensity measurements care must be taken to see that development is really complete, that is, *all the soluble gelatine must be washed away*. The print must be moved about quite vigorously in the flowing water until the latter ceases to be discoloured by the pigment.

After development the frame is unclipped and the back support removed. The remainder, consisting of the waxed transfer paper and the insoluble gelatine film adhering to the single transfer paper that has been shellaced to the metal border, is then removed to the electroscope room and left till next day to dry. It will be found that when the whole is dry the waxed paper can be easily stripped off, leaving the insoluble film adhering firmly to the metal border. The film is then ready for examination with the electroscope.

The reader is advised to consult the makers' pamphlets for a full description of carbon printing. The above outline has been given only to emphasise certain points which are important for the success of these experiments. And the most important point of all we have left till the last. It is this. *From the precision point of view, carbon tissues are quickly spoiled by damp, fog and the fumes that generally occur in the atmosphere of a chemical laboratory. For this reason they should be used as soon after purchase as possible, being kept in the meanwhile in a calcium chloride desiccator.* It is by far the best to do all the printing at once while the tissues are fresh, afterwards examining them with the electroscope at leisure.

Though not so often practised as the silver method of printing, the carbon process will be found to be far more fascinating. With a little practice and ordinary care, it is very little more troublesome to work with success than the commoner process.

Summary.

(1) A simple photometric arrangement is described in which the negative is replaced by a carbon print and the measurements are carried out by means of α-rays and an α-ray electroscope. This radioactivity photometer has been applied to the investigation of the intensity distribution in X-ray crystal photographs.

(2) As a preliminary test the apparatus was used to find the intensity ratio, CuK_α/K_β. In X-ray photographs of the (muscovite) mica cleavage plane it was found that, 1st order : 2nd order : 3rd order = 66 : 31 : 100, and that the intensity ratio CuK_α/K_β is $6 \cdot 2$, when corrected for the absorption of the window of the X-ray tube.

(3) A further test with a double-wedge photograph of the 3rd order K_α mica line gave still more satisfactory intensity values for the seven steps of the wedge.

(4) It is shown how the apparatus may be made to give a curve showing a relation between X-ray intensity and α-ray intensity which is very approximately linear through the origin.

(5) The application of this property to the determination of the integrated intensity of X-ray spectra is discussed and a method outlined.

(6) The method of preparation of the carbon prints for the radioactivity photometer is described.

The above research was carried out in the Davy Faraday Research Laboratory of the Royal Institution, to the Managers of which I am indebted for all facilities. To Sir William Bragg I wish, as always, to express my deep indebtedness for his keen interest and encouragement. I owe my polonium deposits to the late Mr. Harrison Glew and his son, Mr. E. Glew, of 156, Clapham Road, S.W., and to M. M. Ponte, of the École Normale Supérieure, Paris. While studying the best method of preparing the carbon prints, I was assisted throughout by Mr. H. Smith, of the Davy Faraday Laboratory, and for this valuable help, which saved me much time, my best thanks are due to him.

Further Developments of the Method of Obtaining Strong Magnetic Fields.

By P. Kapitza.

(Communicated by Sir E. Rutherford, O.M., P.R.S.—Received June 11, 1927.)

Introduction.

In previous publications* the author has described a method of obtaining strong magnetic fields, and also several experimental researches using these methods.

The general principle of this method is to obtain a field in a coil for a very short time only—about 1/100 second—thus making it possible to use a great amount of electrical energy without heating up the coil. Powers up to several thousand kilowatts are necessary for obtaining strong magnetic fields. Such large powers are easily obtained for a 1/100 second by means of specially constructed accumulator batteries, which make it possible to obtain fields of over 100,000 gauss in a sufficiently large volume to make investigations on such subjects as the Zeeman effect.†

The success of this method has led to the attempt to obtain fields of an order approaching 1,000,000 gauss. A magnetic field of this order requires about 50,000 kilowatts, and this calls for a source of energy about 50–100 times larger than that used in the accumulator experiment. From the experience gained in the previous investigations, it was realised that an accumulator battery would not be practicable, firstly, because such a battery is not very lasting, and, second, because it would be very difficult to have a sharp break of direct-current of such power. A dynamo short-circuited for a small fraction of a second appeared to provide a much more convenient source of energy. In this case the energy is accumulated as kinetic energy in the rotor of the generator, and is converted into electrical power when the machine is short-circuited. It is known from electrical engineering experience that it is possible to obtain large impulses of current in this way.

Rough calculations showed that to give the required power the machine must be of such a size as to give about 2,000 kilowatts at continuous rating. This indicates the scale of the experiment, and, to carry it out, it was necessary to

* 'Roy. Soc. Proc.,' A, vol. 105, p. 691 (1924); 'Roy. Soc. Proc.,' A, vol. 106, p. 602 (1924); 'Proc. Camb. Phil. Soc.,' vol. 21, p. 511 (1923).

† Kapitza and Skinner, 'Roy. Soc. Proc.,' A, vol. 109, p. 227 (1925).

build a special laboratory and to design machinery in a special manner. In this paper it is proposed to give only a brief account of the machinery used, avoiding all technical details. At a future date the writer hopes to give a more detailed description of the apparatus and the principles on which it is designed.

The whole problem falls into five parts, viz. :—

(1) The generator which supplies the large power.
(2) The switch which governs the power.
(3) The timing arrangement.
(4) The coil.
(5) The measurement of the field.

Each of these will be dealt with in a separate section.

The Generator.

An electrical single-phase A.C. generator of the turbo-alternator type was chosen, the main idea being to use half a cycle of the current only. The circuit is made when the generator E.M.F. has its zero value ; the current then rises and drops again to zero, and at this moment the break is made. This method has the following advantages. First, it permits a machine with a high peripheral velocity of the rotor to be used, this type of machine giving the largest powers on short circuit, as can be shown quite generally. Second, by using A.C. current the problem of breaking the power current is made much easier, as the current which is broken is many times smaller than the maximum current ; thirdly, by using a revolving field, sliding brushes or commutators are avoided. Finally, one important advantage remains. In fig. 1 the process of short-circuiting the

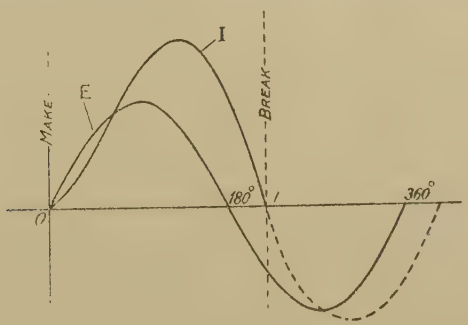


Fig. 1.

machine is represented, Curve 1 representing the current, and Curve E the electro-motive force. During the first half of the period when the electro-motive force builds up the current, it has to overcome the ohmic resistance of the

circuit and create a magnetic field in the coil as well as in the leakage fields of the machine. This is done by taking the energy out of the kinetic energy of the rotor which is gradually slowing down. In the second half of the cycle, when the voltage is reversed, the current still persists, owing to the electro-magnetic energy stored in the coil and in the machine. This energy is partly dissipated in heat owing to the ohmic resistance of the circuit, but a large part of it returns to the rotor, increasing its kinetic energy. A comparison with the case of a D.C. source of energy, such as the accumulators, where we have to short-circuit the coil before breaking the current,* and where the whole of the magnetic energy stored in the magnetic field is dissipated in heat, shows that with an A.C. generator a large amount of this energy goes back to the rotor, the heating up of the coil being some 10 per cent. to 15 per cent. smaller than in the case of accumulators. This makes it possible to obtain stronger magnetic fields with a coil of the same volume and for the same time.

The one difficulty in using A.C. current for our purpose is that the current in the coil has a sinusoidal form, *i.e.*, the magnetic field is never constant. But in some experiments, as in work on the Zeeman effect, a constant field is required for a certain fraction of a second during which the spark exists in the coil.

This difficulty can be overcome by a special design of the excitation winding of the machine. This is ordinarily designed to give an E.M.F. as close as possible to a sine curve, but in our case the winding must be arranged so as to produce a current wave with a flat top. To do this the winding is divided into concentrated coils, one of half pitch and the other of full pitch, the current in each being independently adjusted. These are marked by 1 and 2 respectively in fig. 2,

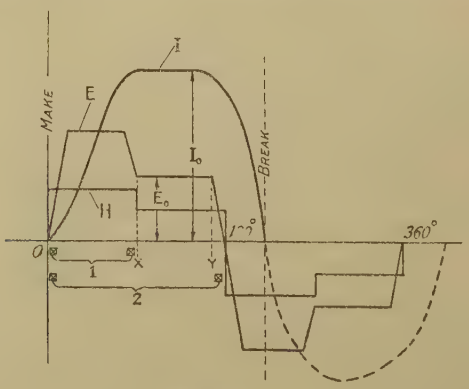


Fig. 2.

* 'Roy. Soc. Proc.,' A, vol. 105, p. 700 (1924).

and the shape of the field is given by the curve H. This field will produce the electromotive force E, and the current I as shown in the same diagram.

To see that the current curve will have in this case a flat top, we consider the general relation

$$E = L\frac{dI}{dt} + RI, \qquad (1)$$

where L and R are the resistance and self-induction of the machine and outside circuit. Now, if we adjust the electromotive force E_0 so that the current satisfies the relation

$$I_0 = \frac{E_0}{R} \qquad (2)$$

at the moment x, then, as can easily be seen from (1), the current will remain constant from the moment x to the moment y. In this way it is possible to obtain a constant magnetic field in the coil for a few thousandths of a second (see fig. 11). Closer consideration shows that the power obtained by the combination of these two fields is slightly smaller than that obtained when using the full pitch winding (2) only, but in most experiments a constant field is only needed when the field is weak. If the field is strong, the experiment can be performed in a much smaller fraction of a second, during which the sinusoidal current, obtained by using the full-pitch winding only, is sufficiently constant.

The general picture of the phenomena which occur in an A.C. generator when short-circuited is well discussed in the Electrical Engineering literature,* and it is well known that the A.C. type of generator at the first moment of short-circuiting gives a very large peak of current which gradually dies out.

This phenomenon enables us to use a much smaller machine. Our generator is designed to have the largest possible value of peak current on short-circuit, and herein lies the essential difference between our machine and a commercial generator, these latter machines being designed, for reasons of safety, to have as small a ratio of transient short-circuit current to normal current as possible. The main principles of the design, therefore, are to have as low an impedance as possible and a strong construction to bear the large short-circuit forces.

The construction of the machine was entrusted to Messrs. Metropolitan Vickers, of Manchester, who took a great interest in the work and carried it out on the most favourable terms. We are greatly indebted to Engineer M. Kostenko, of Petrograd Polytechnical Institute, Prof. Miles Walker, of the

* C. P. Steinmetz, "Theory and Calculation of Transient Electric Phenomena and Oscillations."

University of Manchester, and Engineer G. Kuyser, of Messrs. Metropolitan Vickers, for help and advice. The success of the design of the machine owes much to the wide experience of these specialists and their kind interest in my work. I am chiefly indebted to Prof. Miles Walker, to whom several parts in the construction of the machine are due, especially the ingenious design of the end connection of the stator winding in the copper boxes.

The general view of the machine as it is installed in the laboratory is shown on fig. 3.

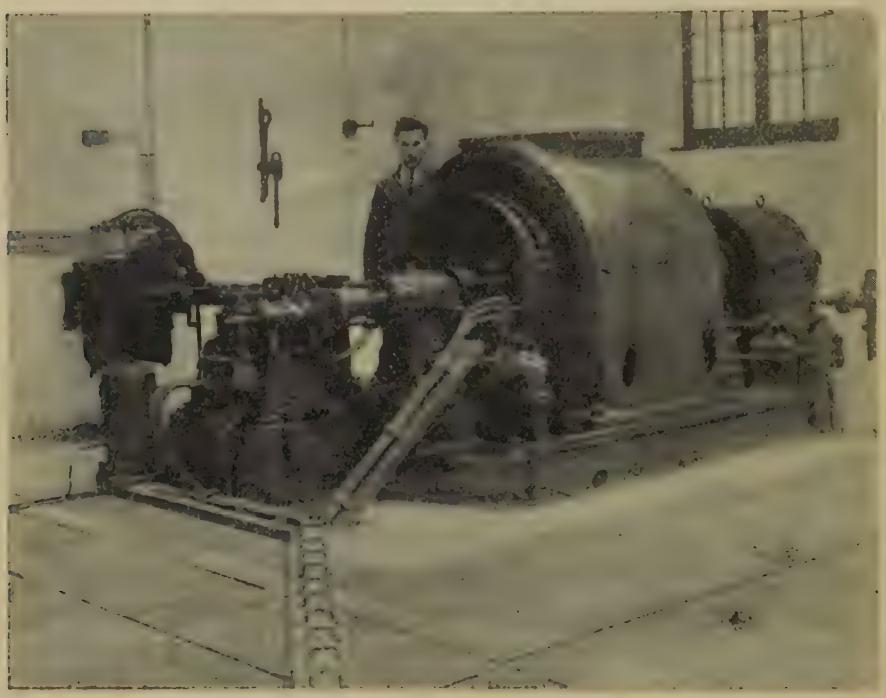


Fig. 3.

As has already been mentioned, the machine is of the normal Turbo-Alternator type, with the exception that certain parts have been strengthened, viz., the stator casting, the bolts and the keying-in of the punchings, etc. This is necessary, as the machine has a very small impedance, and the electro-dynamical forces developed are of such a value that the matter of strengthening these parts is of great importance. The windings, especially the end connections, have been enclosed in special copper-boxes, which not only reinforce them, but also considerably reduce the impedance of the machine.

The steel rotor is 52·1 cm. in diameter, and bears the full pitch and half-

pitch windings. To decrease the impedance a number of damper bars forming a Squirrel cage are embedded in the steel of the rotor. The rotor can make up to 3,500 revolutions per minute, corresponding to a peripheral velocity of 95·7 metres per second; half a cycle is therefore performed in about 1/100 second. The weight of the rotor is $2\frac{1}{2}$ tons, and its moment of inertia 59·6 kilograms/metres^2. The inertia of the rotor is sufficient of itself to provide the necessary amount of kinetic energy, so that no extra fly wheel is required.

The stator is made of punchings cut out of single sheets of Stalloy iron, and these are securely fastened to the steel stator frame. The total weight of the machine is about 13 tons.

As the machine is only run for 15 to 20 minutes during an experiment, there are no special ventilation or cooling arrangements. This considerably simplifies the construction and reduces the windage loss. About 80 h.p. is required to drive the machine, well over-excited, at full speed. The D.C. motor used for spinning up our machine is directly coupled with the rotor and fastened to the same bedplate as the generator.

Both machines and the automatic switch rest on a concrete foundation, 14 ft. 5 ins. long, 7 ft. wide, 3 ft. deep, and about 30 tons in weight. This concrete block rests on a layer of special anti-vibrational material called Mascolite, made of felt and cork. Underneath this felt is another concrete block about $1\frac{1}{2}$ ft. thick, resting on the ground. This arrangement is necessary to absorb the shock which would occur during the short-circuiting of the machine. As the dynamo is further in the ground than any other foundation in the building, any shaking of the building is prevented. The use of such a heavy foundation and of the anti-vibrational material reduces the shock to a very small amount, and in certain experiments even a small shock would be very inconvenient. For instance, in taking spectrograms, a very small shake of the spectrograph would seriously affect the sharpness of the line. To avoid this, the coil in which the experiment is performed is placed about 20 metres away from the machine, the current being supplied to the coil by heavy leads. The disturbance which travels through the ground at a velocity of 2,000—3,000 metres per second does not reach the apparatus until after 1/100 second from the beginning of the short-circuiting, *i.e.*, when the experiment is over. The leads from the machine to the coil consist of six cables—three in parallel to each pole, the cables being interleaved to reduce the self-induction of the rather long connections. As the repulsive forces between the leads are very considerable, the cables are securely fastened together by strong bronze clamps. The loss in the cables is less than 4 per cent. of the total power.

The general arrangement for starting and regulating the speed is very similar to that known by the name of Lennard. Another D.C. generator, driven by an ordinary A.C. motor, is used for supplying the current. The excitation of both the D.C. driving motor and the D.C. dynamo are supplied independently from an accumulator battery. Thus, by means of small rheostats in the excitation circuit the machine can easily be started and kept at any required speed. The use of an accumulator battery for excitation makes the whole arrangement extremely stable.

The currents for the full and half-pitch winding of the rotor of the larger generator are independently supplied by two 6 kilowatt 60 volt D.C. generators. These two small dynamos are excited by the same accumulator battery. In this way the excitation current in the turbo-generator can be adjusted with great accuracy and kept very steady without using heavy rheostats. A further advantage of this arrangement is that during the short-circuiting the possible large variations in the current both of the driving motor and the excitation motor, cannot penetrate into any other laboratory circuit.

The accumulator battery has a rating of 150 ampere hours and 220 volts containing 120 cells.

Before installation in the laboratory the generator was carefully tested at the works. It was first necessary to determine the stresses in different parts of the generator produced by the electrodynamical forces which occurred during the short-circuiting. As the generator has a very small impedance the short-circuit current is much larger than in the usual type of machine and some anxiety was felt as to the safety of some of the more highly-stressed parts. An estimate of the magnitude of these forces required a knowledge of the leakage fields in the air gap. These were measured by a search coil, and it was calculated that the machine would just stand the short-circuiting. For the short-circuiting test a simple oil switch was used, which made the current at some random moment in the phase, so that it was impossible to get the maximum short-circuit current except by chance. Actually, the largest amplitude of the current was 72,000 amperes, and the voltage 2,250, corresponding to a power of 160,000 kilowatts. From this it was calculated that the maximum value of the power of the machine, when short-circuited at zero voltage, would be 220,000 kilowatts. In actual work the machine is short-circuited on a coil whose impedance is approximately equal to that of the machine, so that the total power is only half of the above, and of this one half is lost in the machine, the other half going into the coil and producing a magnetic field. Thus 55,000 kilowatts are available for the production of a magnetic field.

The change in the velocity of the rotor during the short-circuiting is about 10 per cent., so that 20 per cent. of the kinetic energy is drawn out.

The Switch.

The General Theory of the Switch.—From the previous part of this paper the general requirements of the switch are made clear. The time at which the current is made is the moment at which the tension is zero (see fig. 1); it corresponds, therefore, to a definite position of the rotor relative to the stator. The time of the break corresponds to the moment when the current drops to 0. This moment occurs when the rotor has moved more than 180°, and it lies somewhere between 180° and 360°, according to the impedance of the circuit. If the predominant part of the impedance is the ohmic resistance, zero current will occur nearer to 180°, while if the inductance is the predominant part, the zero current will occur nearer to 360°. Thus only the approximate range of the moment of the break can be fixed, which varies with the character of the load and the shape of the rotor fields, the exact moment being determined experimentally each time.

It follows that the switch must work synchronously with the machine in such a way that it makes the current at a fixed moment, but the breaking of the current must be made adjustable within certain limits. Obviously the simplest way of making the switch work synchronously with the generator is to have it operated from the shaft of the dynamo. The switch must be designed in such a way as to suit the requirements of the break and make. The circuit has a large self-induction, so that when contact is made the current rises slowly and the exact time of contact is of no great importance. From previous experience with the accumulator it has been found that it is quite easy to make this time so short that no heat is generated in the brushes of the switch during the make.

The conditions of the break are much more difficult. If we break our circuit very quickly, exactly at the moment when the current is 0, no arcing will occur in theory, but in actual practice it is impossible to make an instantaneous break at this moment. It can be seen that even in a case where the break is 3/10,000 second before or after the current is 0, the current may reach a value of about 3,000 or 6,000 amperes; a current which it is both impossible and dangerous to break instantaneously. The magnetic energy $\frac{1}{2}LI^2$ stored in the dynamo and in the coil has to be dissipated somehow or other, and a sudden break without arcing will transform all this energy into potential energy charging up the capacity of the generator. As this capacity is very

small the potential will rise to very large values, and finally break down the insulation. To avoid this over-tension a condenser is introduced in the circuit during the moment of the break, a capacity of 50 microfarads being sufficient for the purpose.

Fig. 4 shows an oscillogram of the voltage across the break and of the current in the main circuit. It is essential that the brushes of the switch begin to separate (point a) slightly before the moment at which the current drops to

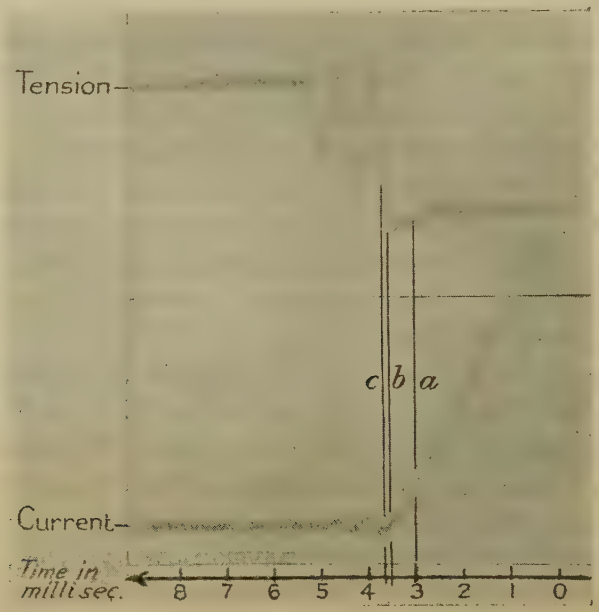


Fig. 4.

zero (point b). On separation, as may be seen by Curve 1, the potential across the break rises to the arcing potential, about 60 volts for two copper contacts in series, and an arc starts between the brushes. The potential rises slowly as the current drops to zero (point b), when the arc is extinguished. As is seen from the oscillogram, the time of duration of the arc is about 0·0005 seconds, and during such a short interval the heat developed does not damage the brushes of the switch in spite of the heavy current. The current then changes its direction, discharges the condenser and recharges it in the opposite direction to approximately double the voltage generated by the machine. These oscillations continue till the condenser is switched out of circuit.

In the interval of time between the points b and c the voltage is below the arcing potential so that no arcing occurs. For the successful operation of the

switch it is important to cool the copper contacts of the break during this interval of time, so as to make any subsequent arcing impossible, and for this purpose, a strong blast of cold air is introduced between the brushes and the plate. The condenser not only prevents over-tension of the machine, but also increases the time from *b* to *c* during which the brushes of the switch may cool.

The Design of the Switch.—The above considerations show that it is important to have a switch operating with an accuracy of 3/10,000 second, relative to the moment when the current drops to zero. In order to ensure this the contacts must be separated by 0·5 mm. in 3/10,000 second, such a motion requiring an acceleration equal to about one thousand times that of the gravitational field. For instance, if the weight of the contact brushes is about 1 kilogram, a force of at least one ton must be applied very suddenly in 3/10,000 second.

A detailed account of the calculations and resulting design of the switch would take up too much space ; I will only give schematically a description of the construction of the switch which has satisfied the above-mentioned requirements.

The general view of the switch is shown in photograph (fig. 5) and a scheme

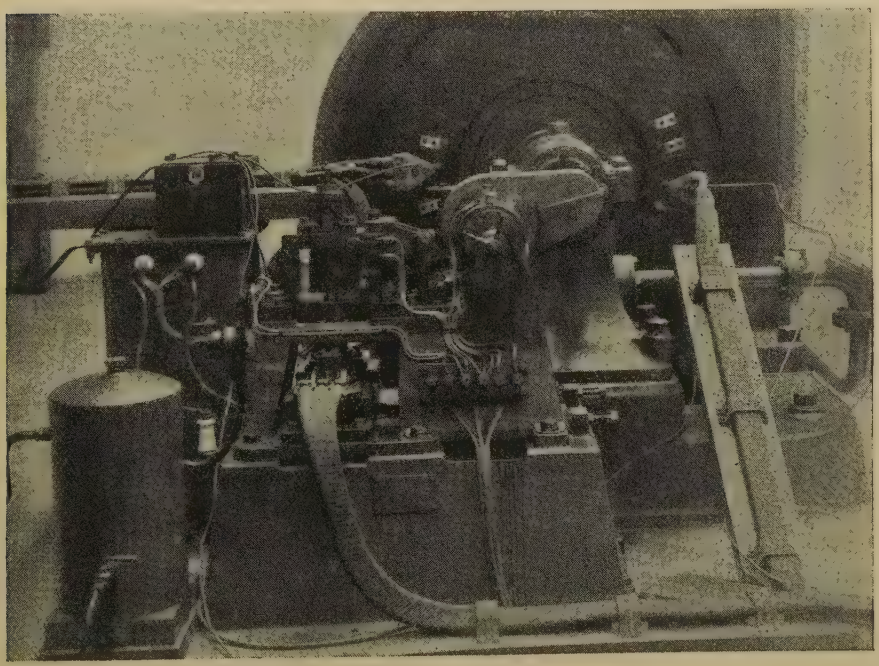


Fig. 5.

of operation in fig. 6. The automatic switch is placed on the same foundation as the generator in such a way that it can be mechanically coupled to the generator shaft. By a silent chain (1) a cam (2) is driven from the shaft of the dynamo at half the angular velocity of the main rotor. By means of a lever this cam (3) pushes the rod (4), which is pressed back to the cam by a strong spring (5). The push rod through a link (6) may press, if engaged, on the joint of the two vertical links (7) and (8); the link (8) is connected to a vertical rod with a copper plate at the end (9), and this copper plate presses on two brushes (10) which form the contacts of the main circuit. The process of operation of the switch is shown schematically on drawing (fig. 6) in successive positions. In fig. 6A of this drawing is seen the position of the parts of the switch just before operation. The cam (2) freely rotates as the lever (3) is prevented by a trigger (11) from falling down on to the cam. The push link (6) is in a downwards position and is not connected with the push rod (4). The links (7) and (8) are on the right-hand side and the copper plate (9) is kept well off the brushes (10) by the spring (12). Now in order to make the switch operate at the right moment, a current is passed through the electro-magnet (13) which pulls down the trigger (11). The lever falls down on to the cam shaft and the push rods begin to make a reciprocating motion as the lever follows the shape of the cam. A current is also sent through the electro-magnet (14) which, by means of the link (15), pulls up the push link (6). At the moment when the lever is on the deepest part of the cam (fig. 6B), the push rods catch in, and in the further motion of the cam the links (7) and (8) begin to straighten up and the copper plate (9) is pressed on to the brushes (10), and the contact is established (fig. 6C). This lasts until the second elevation of the cam comes to the lever (fig. 6D), and this pushes the push rod further on. The links (7) and (8) are thrown to the left-hand side and the copper plate separates from the copper brush. Finally (fig. 6E) the copper plate is well off the copper brush, and as the current is broken in the electro-magnet (13) the lever (3) is again clamped by means of the trigger (11) in such a position that it will not touch the cam. A second small switch (16) is underneath the main switch; the purpose of this small switch is to connect the 50 microfarad condenser (17) between the brushes (10) after the make is effected. This switch (16) is pulled up by a weaker spring (18) and breaks the condenser's circuit only at the moment when the copper plate is well separated from the brushes. In fig. 6 is also shown a nozzle (19), which is connected to a reservoir with compressed air (12 atmos.). By an electrically-operated valve the reservoir is discharged and a heavy blast of air is produced in the air gap of the switch during operation.

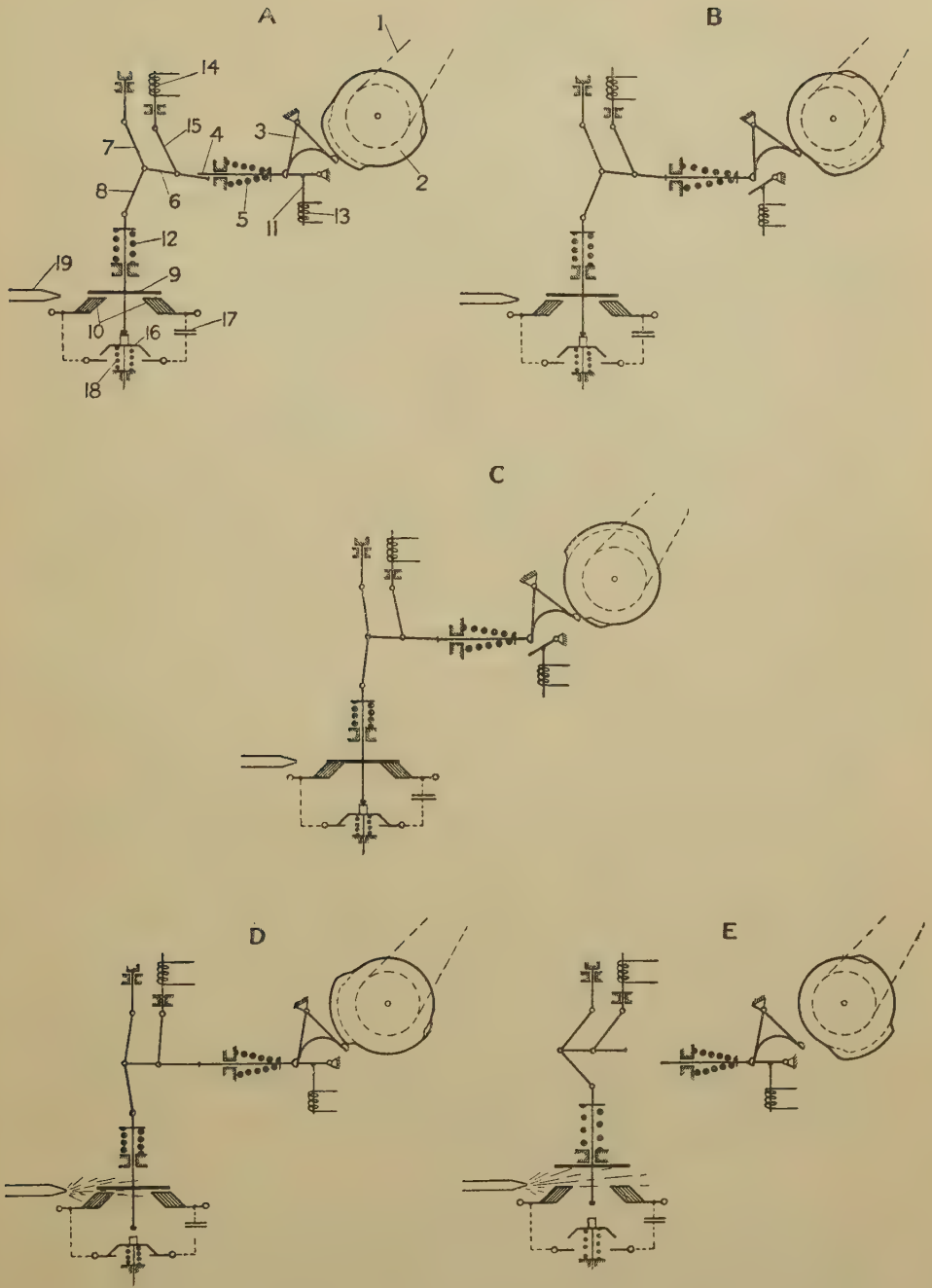


FIG. 6.

It has already been mentioned that only the make has a permanent phase relation with the rotor, the break being adjustable. This is secured by means of the cam (2) being made in two parts, one, having the elevation for the make, being permanently fixed on the shaft, and the other having the elevation for the break, being adjustable within certain limits.

The main features of the design of the switch are as follows. In the first place the switch is not operated by springs, but by forces from the cam which may be made much larger than any obtainable from springs of a small mass ; secondly, the operation of the switch is strictly synchronous with the motion of the rotor ; and thirdly, the roller of the lever (3) is engaged with the cam (2) only for a single revolution. A very great pressure may therefore be allowed between roller and cam without appreciable wearing in a single engagement, large accelerations being thereby attained. The chief difficulty is that large stresses occur from the inertia of the parts themselves. In order to have the various links evenly loaded during the moment of the make and the break, the shape of the cam must be closely calculated and carefully manufactured. It is difficult to keep the safety factor of such a switch above $2\frac{1}{2}$. The switch was made of special high quality steel, and I am indebted to Messrs. Peter Hooker, Ltd., for the very careful and precise manner in which its manufacture was carried out. The switch works well and gives no trouble during the experiments. I hope to give a more detailed description of the method of designing this switch later.

Finally, the successful operation of the switch depends a great deal on the amount by which the phase of zero current varies with such factors as the change in the resistance of the coil owing to its heating up, and to the slowing down of the machine during the moment of the break, these factors varying in different experiments with the same coil. Calculation shows that the moment of the break shifts by 2°, when the number of revolutions changes by 25 per cent., and a similar shift is produced by the change of the resistance due to heating up the experimental coil, made of copper, to 60°. The shift is small but still appreciable, and when the coil is heated up to 125°, considerable care has to be taken in adjusting the moment of the break in accordance with the power which is drawn from the generator.

The Timing.

During the hundredth of a second that the experiment lasts, a number of manipulations have to be timed, and special devices are required for the purpose. First of all the relays of the automatic switch have to be

operated at the proper time. It is important that the relay (13) should release the trigger exactly at the moment when the cam has the position of fig. 6A, when it nearly touches the roller on the lever (3), otherwise under the pressure of the heavy spring (5) the lever (3) may fall from such a distance that it may not only damage but break the cam. The relay (13) has to be set free again before the cam makes a full revolution, to trig the lever again. Relay (14) and the one which operates the valve on the blast arrangement (19) have to be set in motion at the proper time. All these relays are made to operate very quickly, by keeping them small and by sending through them a power of one or two kilowatts. Such a current can be maintained in the relays for only a few seconds without over-heating, so that arrangements must be made to break the current after this time.

On the other hand in the oscillograph room the photographic plate, on which the oscillographic record of the current is taken, has to fall before the images from the oscillograph mirrors exactly at the moment when the machine is short-circuited. All these operations have to be controlled automatically by a timing device. In the present paper we will describe only the general idea underlying the timing.

To start the experiment, a push button switch is pressed and sends a current through a relay which releases the falling plate in the oscillograph room. This circuit has in series a brush and a sliding contact which is coupled with the motion of the cam of the automatic switch, so that the current through the relay can only pass when the rotor of the dynamo is in a definite position. Thus the photographic plate starts to fall synchronously with the machine. In its fall the plate operates a small switch which lets the current into the relays of the automatic switch. These relays have again in series sliding contacts fixed on the cam of the automatic switch, thus permitting the current to operate the switch only when the cam has a definite position. When the plate has finished its fall, it operates another switch which breaks the current in all the relays.

Great care has been taken in designing these timing devices, as on their proper working depends the success of the whole apparatus. As the energy brought into play in each experiment is very large, an accident might have rather serious consequences.

To provide against failure of the automatic switch a heavy high-tension oil switch of the ordinary remote controlled type is introduced into the circuit between the automatic switch and the dynamo. This switch is closed a few minutes before the beginning of the experiment and the button operating the

automatic switch sends current into the releasing device of the oil switch. The switch opens in about one-tenth of a second and sufficient time elapses before opening for the experiment to be performed.

The Coil.

The Theory.—In my previous experiments with accumulators, the electrodynamical forces which occurred in the coil were always marked, but never reached such a value as greatly to affect the shape and the rigidity of the coil. In the present experiments, where a much larger power is brought into play, these forces reach much larger values. When the field inside a copper coil of 1 cm. inside diameter exceeds 200,000 gauss, these forces reach such a value that the coil bursts. The problem of making a sufficiently strong coil to resist magnetic fields above this value was probably the most difficult part of the whole experiment, and a great deal of research was necessary to find a construction for the coil which would stand the strain.

The first part of the problem was a theoretical study of the strains produced in the coil when a current is flowing, and of how by a proper choice of the cross-section of the coil these strains may be reduced and brought to such a form that reinforcement of the coil is made easy. The second part of the problem was more mechanical, consisting of working out the actual construction of the coil and of the reinforcements.

Let us first deal with the theoretical part of the problem. On fig. 7 is shown

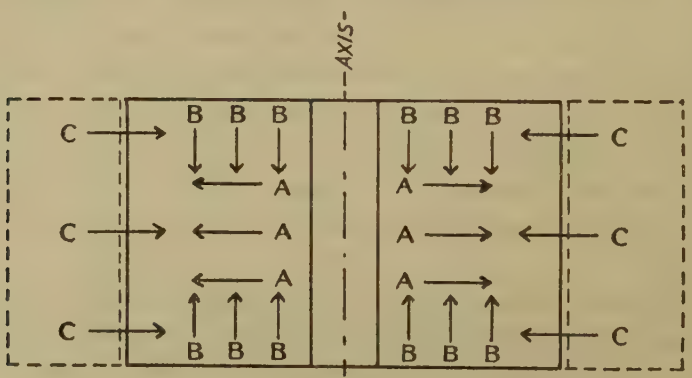


Fig. 7.

the cross section of a circular coil. When a current is flowing, electrodynamical forces will act in the body of the coil. The principal forces, shown by area A, are radial and tend to increase the diameter of the coil. These are the forces which tend to burst the coil.

If we reinforce the coil by some kind of bandage placed round it, as shown by a dotted line in fig. 7, a surface force, shown by area C, will be produced and may prevent the bursting. This arrangement will work within certain limits, beyond which the copper will simply flow to the side in the direction of the axis of the coil. To prevent the deformation by reinforcement from the side is technically very difficult, if not impossible. There is, however, no actual need for this side reinforcement, as nature comes to our help by another set of electrodynamical forces marked by area B in fig. 7, which act so as to stop the copper flowing to the sides. Clearly the problem is reduced to finding a coil of such a construction that there is a certain balance between the forces A, B, and C. The ideal case would occur if we could reduce all these forces to a uniform (hydrostatic) pressure, so that there will be no shearing strain in the copper of the coil, and, if the reinforcing ring is made sufficiently strong, it will keep its shape without any side supports. It may be theoretically possible to find a shape of coil satisfying the above-mentioned requirements, but such a coil will not be of a square cross-section, and in practice it is very difficult to wind a coil that is not rectangular in section. The next approximation is to make a coil with a stepped cross-section, in which case it is possible to reduce the shearing stresses in the coil if the proper ratio between the length of the coil and the outside and inside diameters is chosen.

These forces can be calculated and measured in the following way. Let us assume that a current of uniform density I is circulating in the coil (see fig. 8), and imagine a coaxial ring placed somewhere in the body of the coil. If we choose cylindrical co-ordinate z along the axis of the coil, and r the radius,

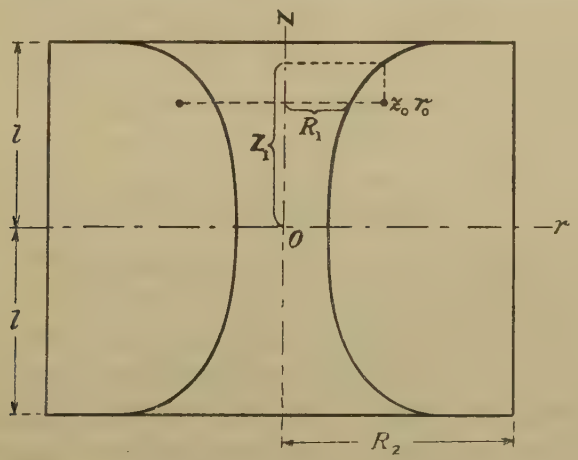


Fig. 8.

then we may determine the position of the ring by z and r. Now, the magnetic flux from this coil through the ring is $\phi = \mathrm{ISM}$, where M is the coefficient of mutual induction of the coil and the ring and S is the surface of the cross-section of the coil. We may regard ϕ and M as continuous functions of z and r. Then the two components of the magnetic field at any point will be :—

$$\mathrm{H}_z = \frac{1}{2\pi r}\frac{\partial \phi}{\partial r} = \frac{\mathrm{IS}}{2\pi r}\frac{\partial \mathrm{M}}{\partial r}, \tag{3}$$

$$\mathrm{H}_r = \frac{1}{2\pi r}\frac{\partial \phi}{\partial z} = -\frac{\mathrm{IS}}{2\pi r}\frac{\partial \mathrm{M}}{\partial z}. \tag{4}$$

The mechanical force on unit of the volume of the coil in the directions corresponding to r and z will be

$$\mathrm{F}_r = \mathrm{IH}_z = \frac{\mathrm{I^2S}}{2\pi r}\frac{\partial \mathrm{M}}{\partial r}, \tag{5}$$

$$\mathrm{F}_z = -\mathrm{IH}_r = \frac{\mathrm{I^2S}}{2\pi r}\frac{\partial \mathrm{M}}{\partial z}. \tag{6}$$

We see that M is in some respects similar to a potential to the force F.

At a point with co-ordinates z_0 and r_0, the copper will be compressed in the r direction with a force

$$\bar{\mathrm{F}}_{r_0} = \frac{1}{r_0}\int_{\mathrm{R}_1}^{r_0} r\mathrm{F}_r\, dr, \tag{7}$$

and in the z direction

$$\bar{\mathrm{F}}_{z_0} = \int_{\mathrm{Z}_1}^{z_0} \mathrm{F}_z\, dz, \tag{8}$$

where R_1 is the radius of the inside surface of the coil with the co-ordinate z_0 and Z_1 is the Z co-ordinate of the same surface with the radius r_0. The integration is easily carried out by means of (5) and (6), giving

$$\bar{\mathrm{F}}_{r_0} = \frac{\mathrm{I^2S}}{2\pi r_0}[\mathrm{M}(r_0 z_0) - \mathrm{M}(\mathrm{R}_1 z_0)] \tag{9}$$

and

$$\bar{\mathrm{F}}_{z_0} = \frac{\mathrm{I^2S}}{2\pi r_0}[\mathrm{M}(r_0 z_0) - \mathrm{M}(r_0 \mathrm{Z}_1)]. \tag{10}$$

For the copper to be uniformly compressed we must make $\bar{\mathrm{F}}_r = \bar{\mathrm{F}}_z$ everywhere in the coil. Expressions (9) and (10) show that this can only be obtained if the value of M on the inside surface of the coil is constant ($\mathrm{M}(r_0\mathrm{Z}_1) = \mathrm{M}(\mathrm{R}_1 z_0)$). But it is easily seen that the surface along which M is constant will coincide with the magnetic lines of force. *To reduce the forces everywhere to uniform compression, a coil must have a cross-section of such a shape that the inside unsupported surface coincides with the magnetic lines of force.* If the coil is

not this shape, it is easily seen that the maximum shearing forces (N) which occur in the coil at the point $(r_0 z_0)$ are

$$N = \bar{F}_{r_0} - \bar{F}_{z_0} = \frac{I^2 S}{2\pi} \frac{M(R_1 z_0) - M(r_0 Z_1)}{r_0}. \qquad (11)$$

From this expression we see that *to determine the shearing stress in the coil it is sufficient to know the numerical value of* M *on the surface of the coil only.*

Expression (11) enables us to determine the point at which N reaches its maximum value, and this is the point at which the coil is most likely to give way. For example, in a coil which is not very flat, the point of maximum stress is at the corners which touch the reinforcing winding. In stepped coils, it lies at the inside corner of the steps.

To determine the shearing force N we have to calculate M. From the expressions (3) and (4) and the field equation curl $H = 4\pi I$, we obtain for M the following partial differential equation* :—

$$\frac{1}{r}\frac{\partial^2 M}{\partial r} + \frac{1}{r^2}\frac{\partial M}{\partial r} + \frac{1}{r}\frac{\partial^2 M}{\partial z^2} = \begin{cases} 0 \text{ outside the coil} \\ 8\pi^2/S \text{ inside the coil} \end{cases}. \qquad (12)$$

Butterworth† has shown that a solution of this equation can be obtained in terms of series derived from Bessel functions, and these series were used to calculate the coefficients of mutual induction of solid coils. Similar series for the mutual inductance of coils of rectangular section and coaxial circles have been developed and used to compile tables giving numerical results for the different shapes of coils likely to be used in practice. With these tables it is possible to investigate the effect of varying the shape of cross-section on the maximum shearing stress in the coil. It will probably be difficult to design a coil in which the shearing force is absolutely zero, but by a certain amount of guess-work and successive approximation it is possible to obtain a coil in which the shearing forces are small for all practical purposes.

It is easy to measure the coefficient M on the surface of an actual coil by means of a ballistic galvanometer. This is done by placing at the point of the surface where we wish to measure M, a small coaxial coil made of a few strands of fine wire and connecting it to a ballistic galvanometer. If the current in the main coil is suddenly stopped, the throw of the ballistic galvanometer is exactly proportional to M. The absolute value of M can easily be found from the calibration of the galvanometer by means of a standard mutual induction introduced in the circuit in the usual way.

* Clerk-Maxwell, vol. 2, para. 703.

† 'Phil. Mag.,' vol. 29, p. 572 (1915).

Finally we have to find the force on the bandage, necessary to make it sufficiently strong. The force which presses the bandage ring out is obtained from (9) and is found to be

$$P = \frac{I^2 S}{\pi}\left[\int_{-l}^{+l} M(R_2 z)\, dz - \int_{-l}^{+l} M(R_1 z)\, .\, dz\right]. \qquad (13)$$

Where R_2 and R_1 are the outside and the inside radii of the coil and l is half the height of the coil. As both integrals are taken on the surface they can easily be calculated from the data from the measurement with the ballistic galvanometer or from the tables.

The Construction of the Coil.—In designing the coil we first choose the cross-section and the volume. The cross-section is not only decided by the condition of rigidity, but also by the requirements of the experiment—the degree of uniformity of the field and the space in which the field is to be produced. The efficiency of the coil has also to be taken into account. The volume of the coil is obtained from the rise in temperature which may be allowed in a particular set of experiments. The temperature must never rise above 150°, as this is the maximum temperature which the insulation can safely stand.

It is exceedingly difficult to satisfy all these requirements, and it would occupy too much space to describe the various methods by which the solution of this problem was attempted. By means of specially calculated tables* and graphs, however, it has been found possible to find a coil which will give the required results.

When the shape has been chosen the reinforcement is easily designed by using expression (13). The number of turns is chosen so that the impedance of the coil is approximately equal to that of the generator with the cables.

The cross-section of the strip must not be more than 5 mm. in width, as otherwise the penetration of the current in 1/100 second will not be complete; the resistance of the coil will therefore increase and the copper will not be uniformly compressed. When a coil is properly designed it can be seen that, in spite of the fact that forces inside the body reach very large values (several tons per square cm.), it is possible to compress the material so uniformly that if, instead of copper, we use an alloy of copper with 2 per cent. of cadmium,† which has 90 per cent. of the conductivity of pure copper, and when hardened has about

* Mr. J. D. Cockcroft has drawn up special tables for calculating M, and has worked out several formulæ which are very useful in designing the coils. These will be published at a later date.

† The cadmium copper strip was specially prepared and rolled for us by Messrs. Thomas Bolton, to whom we are much indebted.

four times its tensile strength, then a coil can be made which will stand well over one million gauss without changing its shape. We may expect, of course, that the insulation will give way sooner or later, but at present, in actual practice, the chief difficulty arises from another source.

The coil has to be wound in several layers, so that the following phenomena occur when the current is passing through the coil. The outside layer of the coil first becomes slightly compressed and increases in diameter. The diameter of the second layer then experiences a larger increase consisting of its own increase plus the increase of the outside layer. This will result in a slight unwinding of each layer proceeding from the outside to the inside and gradually accumulating. When the inside layer, which is connected to the other end of the lead, is reached, it may unwind by as much as 1 cm., and the conductor may then be strained to such an extent that it breaks. About five coils failed owing to the middle conductor breaking in this way.

At first an attempt was made to stop this unwinding by making the coil as tight and rigid as possible, the coil and steel reinforcing band being both wound under very strong tension. After winding, the coil was impregnated under pressure with Bakelite varnish and baked, according to the usual Bakelising practice. In this way all the gaps inside the coil were filled with solid Bakelite. As a result of this process the unwinding did not occur until higher fields were reached, but it was still impossible to make a coil which would stand a magnetic field of more than 230-250 thousand gauss. This means that the forces which are brought into play in the coil when the current is flowing are so much larger than those which are applied during the winding process, that it is impossible to make a rigid coil. The only way in which it is possible to make a coil which will stand the strong magnetic field is consequently to allow the unwinding to go on without breaking the winding. This involves introducing a slipping contact in some place in the winding. It is rather an unusual technical construction to make a small slipping contact to stand a current of 30,000 amperes, but as the current is only passed through this contact for a very short time it appeared to be possible. The drawing in fig. 9 shows a coil with such a slipping contact, which is now used in the experiments. The winding (1) is a cadmium copper strip and starts from a manganese bronze case (2) to which it is soldered. The winding is made in four layers, and to the last turn of the winding is soldered a copper ring (3) with a small slot which forms the end of the winding. Inside the coil is placed a cadmium copper cylinder (4) with a helical cut in the middle which forms the inside turn of the winding. The left end of this cylinder has one long slot which joins the helical cut, and many shorter axial cuts. It

fits the ring (3) and is pressed to the side of the ring by means of a conical fibre bush (5). This arrangement makes the sliding contact. The ring (3),

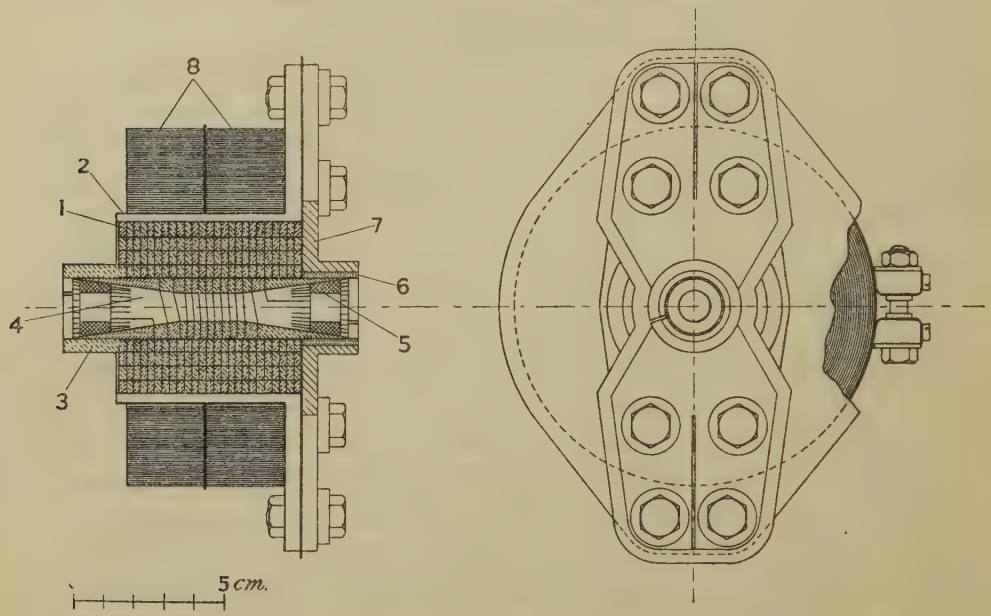


Fig. 9.

connected to the last turn of the coil, is made to slip on the end of the cylinder (4), and thus no stress is produced in the strip. On the other end of the cylinder is an exactly similar arrangement with a ring (6) fastened to a phosphor bronze plate (7) which forms the other lead of the coil. The coil is reinforced from outside by a steel band (8) which is tightly wound on the bronze case (2). The steel used is of a high quality and is hardened and tempered. The forces which attempt to burst the steel ring are large, and reach, for instance, the value of 100 tons when the field inside the coil is about 400 kilogauss.

Special care is taken with the sliding contact arrangement, since it might happen that when the ring (3) moves it would get short-circuited. This is prevented by packing some insulation over the slot in the sliding contact. Special care is also taken with plate (7), as the current is flowing along it in the outside field produced by the coil and the electro-dynamical forces on this plate may reach a very high value; the plate is therefore made sufficiently thick, and is well bolted.

This coil has at present stood successfully up to 320 kilogauss without any sign of damage. The first time this field was obtained the ring (4) of the sliding

contact moved by about 5 mm., but on repetition of the experiment no motion was observed. Theoretically, it can stand a much larger strain, corresponding to a larger field. The question as to what will happen to the insulation remains open. That used in the coil is leatheroid and mica, and the coil is well impregnated with Bakelite and is baked. Will it stand greater pressure? Fortunately there are indications that good insulation in the coil is not necessary, as the magnetic field, which penetrates the coil, prevents any arcing in the body of the coil.

The Method of Measuring the Field.

To determine the field in my previous experiments the current through the coil was measured by means of an oscillograph, and knowing the constant of the coil it was easy to calculate the magnetic field. It is possible to calculate the constant of the coil, but as the coils are not particularly accurately wound, direct experimental methods of determining the constant have been worked out.* These methods cannot be used in my present experiment as a steel band is placed round the coil. This steel band may have a marked influence on the determination of the constant of the coil when a small current is used. With the larger current which is passing in our experiment the steel band is very soon saturated and contributes a constant value to the magnetic field inside the coil. These considerations led to the development of a method whereby strong magnetic fields can be measured directly. The standard method of a search coil connected to a ballistic galvanometer was adopted for our purpose. The difficulty in the application of this method lies in the fact that if we simply introduce a search coil in the magnetic field, the total change of the flux during the experiment will be zero, and no deflection will occur. To adapt this method for our purpose a special automatic switch is arranged so that it short-circuits the search coil which is connected to the ballistic galvanometer. Thus, the current induced in the search coil does not flow in the ballistic galvanometer, but when, somewhere in the middle of the current wave, the break opens the circuit, the induced current is thrown into the ballistic galvanometer. Obviously, the deflection of the ballistic galvanometer will be proportional to the magnetic flux which penetrates the search coil at the moment when the circuit is broken. By means of automatic timing devices a little switch can be made to operate at any chosen moment of the current wave. To mark the moment of operation a small shutter is connected to the switch, so arranged that it blots out the light which falls on the oscillograph exactly at the moment when

* 'Roy. Soc. Proc.,' A, vol. 105, p. 707 (1924); 'Roy. Soc. Proc.,' A, vol. 109, p. 232 (1925).

the circuit of the search coil is broken. In fig. 10 an oscillogram of a current curve is given on which such marks produced by the shutter can be seen on the top of the current wave. The search coil used in this experiment was wound on a well-ground glass cylinder, 6 mm. in diameter, and consisted of 19 turns of closely-wound enamelled wire, gauge 44. The effective area of this coil was calculated from direct measurement by a micrometer. The ballistic galvanometer was calibrated by means of a standard mutual inductance, permanently introduced into the circuit. The calibration was done by a standard ammeter. The current oscillograph was calibrated by means of the same ammeter; in this way the constant of the coil was determined relative to the absolute value of the mutual inductance, independently of the absolute accuracy of the ammeter.

Great care was taken to use such a bifilar arrangement of the leads connecting the search coil with the ballistic galvanometer that no E.M.F. was induced in the circuit from these leads.

A control experiment was made in which the coil was short-circuited during the whole experiment; no deflection of the galvanometer was obtained, showing that no E.M.F. was induced in the connecting leads between the coil and the ballistic galvanometer. The current in the large coil was measured by an oscillograph connected to a shunt, in the usual way, as described in the previous paper.*

From a series of experiments carried out with various field strengths ranging from 45 to 250 kilogauss, it was found that the steel ring gives a constant contribution of about 3,800 gauss, the steel ring being well saturated when the field inside the coil is below 45,000 gauss. Thus, to determine the field inside

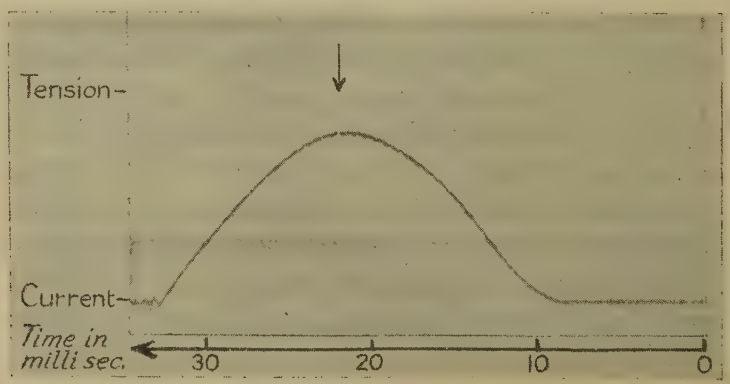


Fig. 10.

* 'Roy. Soc. Proc.,' A, vol. 105, p. 701 (1924).

the coil it is necessary to multiply the current taken from the oscillogram by the constant of the coil (the coil shown in fig. 9 has a constant of 23·2) and to the value obtained add 3,800 gauss. In fig. 10 the current at the moment marked reached the value of 10,760 amperes, and the field inside the coil was 253·7 kilogauss. Agreement between different readings was within 1 per cent. which is quite sufficient for our present research. There are, however, no obvious reasons why the accuracy cannot be increased, if necessary.

The Experiments.

From this general description it can be seen that practically all the apparatus has to be specially made and designed for our experiment. During the experiment three men at least are required for the manipulation of this apparatus :— one watching the machine and the switch, one attending to the oscillograph, and another adjusting the speed and the excitation of the generator, operating the switches and manipulating all the buttons. By means of electrical signals each one of the workers announces that he is ready and then the button is pressed and all the experiment is done automatically. It took some time to make the various arrangements work smoothly and reliably.

It must be remembered that the power brought into play in each experiment is very large. For instance, the energy released in a single experiment approaches that of a field gun. The amount of the energy involved in an experiment is very easily felt from the explosions which occur when the coils burst. Great care must be taken to avoid accidents.

Both forms of the current wave have been used in the experiments. The first is approximately sinusoidal and has been shown in fig. 10. The second, shown in fig. 11, has a flat top and is obtained by using the two excitation windings. From both these oscillograms it will be seen that ripples are present in the current wave, these being produced by the teeth and slots of the machine. The amplitude of these ripples diminishes with increasing saturation as can be seen from the fact that the ripples are smaller in oscillogram 10 where the flux density is higher than in oscillogram 11. In the present experiment these ripples are of no importance, but if required they can be eliminated by the use of condensers.

In oscillogram 11 the time scale is introduced by means of a synchronous motor and interrupter in the light beam. The potential across the break is shown on both oscillograms by a separate curve.

At present, using the coil described in one of the previous paragraphs, and shown on fig. 9, the maximum fields obtained were 320 kilogauss, and in this

case the machine was run at 1,500 r.p.m., about half its maximum speed, only about one fifth of the available power being used. By using the full power we

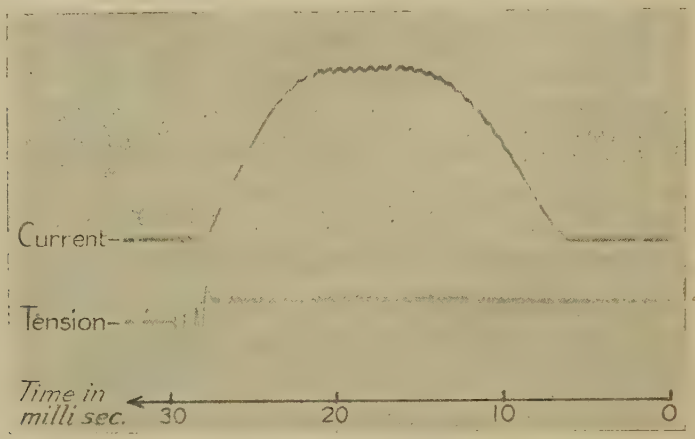


Fig. 11.

may expect to obtain fields up to 700,000 gauss with the coil; and, if we diminish the inside opening of the coil from one cm. to half cm., still leaving sufficient space for several experiments, it is hoped to obtain fields up to 900 kilogauss. The present coil shows no sign of giving way, although many heavy discharges have passed through it, but as the force inside the coil reaches a value for which no experimental data exists, it is very difficult to be certain when the limit of the coil will be reached. The power at present used is ten times larger and the fields three times greater than was used in my experiments with the accumulators. As we advance further, still larger powers must be used; this will involve special precautions in experimenting, gradually increasing the power step by step. So far experiments have been confined to magnetic fields not greater than 350,000 gauss. Experiments have just been made on the change of resistance of bismuth in magnetic fields up to this limit. In a field of about 300,000 gauss, at a temperature of liquid air the resistance of bismuth increases 1000 times, and 50 times at a normal temperature. These changes of resistance are easily observed by oscillographic methods, and it is hoped to publish a paper on this subject shortly.

In order to carry out these experiments, a special Laboratory has been arranged and a set of machines of special design have been constructed. This work has only been possible by the continued support of the Department of Scientific and Industrial Research, who have defrayed the whole cost of the

installation and of the experiments. We are much indebted also to Sir William Pope for providing an admirable Laboratory in which to carry out the experiments.

In the design and installation of the apparatus, many difficult technical and practical problems were encountered, and the successful accomplishment of the work owes much to the help so freely afforded to me by many scientific men and by the firms who constructed the apparatus. In addition to those already mentioned, I am especially indebted for the advice and help given by Mr. H. T. Tizard, Mr. F. E. Smith, Mr. H. J. Thomson, Mr. A. P. M. Fleming, Mr. G. McKerrow, and many others.

My personal thanks are due to Mr. J. D. Cockcroft for much valuable help given me in testing and installing the large generator and in special calculations on the machine and coil, and to Mr. E. Laurmann for his assistance in the conduct of the experiments and in the design of special apparatus.

I would most particularly like to thank Sir Ernest Rutherford for his help and for the very keen interest he has shown during the whole course of the work, and for his appreciation and support of the ideas involved.

Summary.

A further development of the method for obtaining strong fields for a short time is described. This development is obtained by the use of larger powers.

A dynamo and the switch are described, which give very large impulses of current.

Coils able to stand strong magnetic fields are described, and with these coils magnetic fields up to 320,000 gauss in a volume of 2 c.cm. were obtained.

The direct method of measuring magnetic fields of short duration with ballistic galvanometers is given.

The Combination of Nitrogen and Hydrogen Activated by Electrons.

By A. Caress and E. K. Rideal.

(Communicated by T. M. Lowry, F.R.S.—Received April 29, 1927.)

Introduction.

Although the chemical reactions of active molecules have been used to fix their critical potentials, little attention has been paid to the converse process of analysing the mechanism of a reaction by studying the rates of reaction of molecules activated to the energy-levels already known from physical measurements. The method is analogous to photochemistry,* with the advantages that a stream of electrons with a nearly uniform velocity is easier to obtain than quasi-monochromatic light, and that the mechanism is a more general one. The following is an account of experiments on the rate of production of ammonia from nitrogen and hydrogen as a function of the energy of thermions used to activate molecules and atoms. No attempt was made to measure the critical energy increments with great exactness, since that can be accomplished more easily by physical means† ; the object in the present case was to recognise the critical potentials of significance in the reaction.

Previous Work.

Heidemann‡ described the production of ammonia even at the lowest voltages, but subsequent work by Andersen§ and Storch and Olson|| did not confirm this. They detected no combination until the molecular ionisation potential of N_2 (circa 17 V) was reached, after which the reaction rate increased abruptly every 4–7 V. The mechanism proposed was that H_2^+ and N_2^+ appearing at 16 V and 17 V respectively gave H and N atoms on collision, and that increased combination was due to the activation of H by 4 V electrons. Later Kwei¶ found that the NH_3 band spectrum was not excited in hydrogen and nitrogen mixtures until 23 V was reached. This voltage corresponds to the

* We might suggest kinelectron chemistry, or simply kineto-chemistry, since α-particles are also used (Lind).

† See Compton and Mohler for a useful summary. 'National Research Council Bulletin,' vol. 48 (September, 1924).

‡ 'Chem. Z.,' vol. 45, p. 1073 (1921).

§ 'Z. f. Physik,' vol. 10, p. 64 (1923).

|| 'J. Am. Chem. Soc.,' vol. 45, p. 1605 (1923).

¶ 'Phys. Rev.,' vol. 26, p. 537 (1925).

second jump in Storch and Olson's curve. In a subsequent note Olson* explained the failure of Kwei to detect ammonia at 17 V by postulating that NH_3^+ must be present for the spectrum to appear. Thus at 17 V the reactions were considered to be $N_2^+ + e \longrightarrow N'_2$, $N'_2 + N_2 \longrightarrow N_2 + 2N$, the nitrogen atoms then combining with H_2 or H produced by the reaction $N_2 + H_2 \longrightarrow 2N + 2H$; while at 23 V the voltage at which N^+ begins to appear, NH_3^+ is obtained in the same way.

It may be noted that—

(1) The question of whether ammonia can be produced under any circumstances below 16–17 V, where arc formation begins, was left undecided.

(2) Storch and Olson used a diode with a tungsten filament as source of electrons. The two chief disadvantages of this arrangement are (*a*) Below the first ionising potential most of the drop in potential is near the anode. Hence electrons which collide before reaching the anode have much less energy than that corresponding to the applied voltage. If the pressure is such that the mean free path of the electrons between effective collisions is much greater than the distance between electrodes, the amount of activation is again lessened. There is thus only a small range of pressure in which activation can take place on a scale to produce a measurable reaction rate. (*b*) When an arc has actually been set up owing to the neutralisation of the negative space charge by positive ions, there are at least two complications of importance. First, owing to the lower mobility of the positive ions, the potential curve may have a maximum, with the result that electrons fall through a potential larger than the applied voltage. Again according to Compton† the primary quality of an arc is the presence of positive ions. But in addition electrons which have lost their energy fall back into the arc producing a high rate of recombination and intense radiation; these secondary properties of the arc may profoundly modify the reaction.

(3) Storch and Olson noticed a higher rate of combination with thinner filaments and suggested that this was due to the decomposition of ammonia by the filament. The much more likely explanation, that hydrogen atoms, formed by thermal dissociation of molecular hydrogen at the filament, also contribute to the reaction, is discussed below.

* 'J. Am. Chem. Soc.,' vol. 48, p. 1298 (1926).

† 'Phys. Rev.,' vol. 21, p. 207 (1923).

Experimental.

The apparatus employed is shown diagrammatically in fig. 1. A triode was used and the accelerating potential was applied between the grid and the

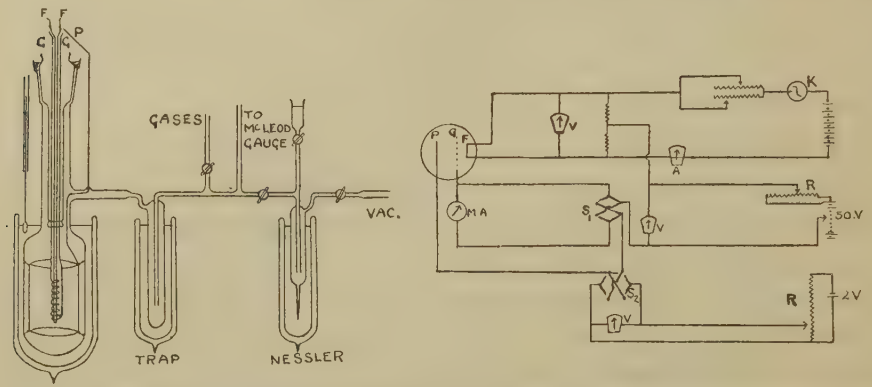


Fig. 1.—On the left is a sketch of the triode shown again on the right with the electrical connections (simplified). F = filament, G = grid, P = plate, V = voltmeters, A = ammeter, R = radiopotentiometers, M.A. = microammeter, VAC = to mercury diffusion pumps, drying tubes, and Hyvac oil pump.

junction of two high resistances (circa 1000 Ω) in parallel with the filament, which was heated by accumulators in series with two rheostats in parallel. The high resistances were adjusted in preliminary experiments so as to be at the same potential as the centre of the filament by the Wheatstone bridge method. A small potential ($< 0{\cdot}2$ V) was applied between grid and plate to collect stray electrons. Fine adjustment of both accelerating potentials was secured by tapping off from radiopotentiometers across a single cell. The switch S_1 made it possible to read either the total emission or the grid-anode current. By means of S_2 the grid-anode potential could be reversed. By a further switch (not shown) the heating current could be passed through both filament and grid-wire in baking out.

Tubes of many types were made up and used in preliminary experiments; the design finally adopted is shown in the figure. The reaction vessel consisted of a cylindrical bulb just large enough to contain the platinum or nickel cylinder anode, capable of total immersion in liquid air. The long narrow tube leading from the bulb to the ground joint served the double purpose of keeping the joint, which had to be lightly greased and mercury-sealed, at a distance from the bulb and of reducing to a minimum the error due to change of liquid air level. The filament and grid were removable by means of the ground-joint. Two tubes were employed, in each case the tungsten filament of diameter 0·2 mm.

and 30 mm. long was stretched between either a spiral molybdenum wire grid of spiral radius 5·5 mm. or a perforated nickel cylinder of radius 3 mm. and holes 1 mm. wide. The anode in the former case was a platinum cylinder of radius 16 mm. and 40 mm. long, and in the latter a nickel cylinder 21 mm. radius and 40 mm. long.

Procedure.

The separate gases were bubbled through sulphuric acid and passed over red hot copper mixed with palladium and then through phosphoric anhydride drying tubes. Mixing was effected in burettes over mercury at atmospheric pressure. Samples of the pure gases and various mixtures were then kept in bulbs lined with phosphoric anhydride from which they could be pumped through the apparatus by way of a mercury-vapour trap kept in liquid air. A further trap, separated by stopcocks from the reaction vessel and McLeod gauge on one side and the pumping system on the other, was used for carrying out Nessler tests on the reaction products. After an experiment the Nessler tube was kept in liquid air and the gases pumped through it. The tube was then isolated and Nessler solution was run in as the tube warmed up. Ammonia was detected by the appearance of an orange colour at the bottom of the inlet tube, where the gas met the solution. After each test, the narrow tube at the bottom of the trap was broken and the tube was washed out with water, alcohol, ether, dry air and finally dried in a flame.

Between runs the reaction vessel was heated in a furnace to about 400° C. and exhausted, the filament being raised to a temperature much higher than those used in experiments, and a high potential being applied between filament and grid and between grid and anode. The baking was continued for a period of from one to six hours and lead tubes were dried as far as possible by a flame. After this the filament could be kept lit at experimental temperatures without increase of pressure, so long as the tube was surrounded by liquid air.

During runs, liquid air was kept at a constant level round the trap and also round the reaction vessel, the level being above the constriction in the tube. By means of a turn-table and a float with a glass pointer kept at a mark in a fixed tube, the level was kept constant to a few millimetres. This was found to be necessary, since molecular hydrogen and nitrogen are adsorbed on surfaces at liquid air temperatures, and the rate at which hydrogen atoms and ammonia are frozen out is also fairly sensitive to liquid-air level.

Before beginning readings, it was always necessary to wait until, first, the adsorption of gases had come to equilibrium, and, finally, temperature equilibrium was attained with the filament lit.

Results.

In preliminary experiments, the procedure of Storch and Olson was adopted, with the modification that a grid was introduced.

A mixture of 75 per cent. H_2 and 25 per cent. N_2 having been pumped in to a convenient pressure, the emission was kept constant by slight variations of the heating current, and the rate of pressure fall was read on the McLeod gauge for various voltages, and finally with no applied voltage. After baking out, a new specimen of the same mixture was introduced and the experiment repeated.

The general results of numerous experiments of this type were—(*a*) At the emission used, a high rate of pressure decrease (nearly half the maximum) was observed with no applied voltage. (*b*) All the experiments showed abrupt increases of rate at about 13 V and 17 V. (*c*) No further jumps were observed up to 50 V; in fact, in some experiments, the rate decreased after 25 V. In general, the results obtained by this method were not repeatable with sufficient accuracy to detect jumps other than the two mentioned. The pressure drop, with no applied voltage, was surprising, although confirming Heidemann's results. Since it was not observed by subsequent workers, a critical examination was necessary. The possible explanations are :—

The presence of oxygen or gases from the tap-grease as impurities. A tube was made up without a ground-joint, and no appreciable difference was observed in the effects. This ruled out impurities with the exception of oxygen. The gases were therefore subjected to a further purification by keeping them for some time in a vessel containing a long coil of glowing tungsten wire. They were passed through phosphoric anhydride drying tubes and a liquid-air trap directly into the reaction vessel. Leaks were also very carefully tested for. These precautions made no difference.

Again, the effect might be due to freezing out of activated hydrogen or nitrogen or both. Storch and Olson found that their blank tests with the pure gases gave negligible rates of fall of pressure. On the other hand, Hughes,* using an oxide-coated platinum emitter, found that hydrogen and nitrogen were frozen out in liquid air after collision with electrons. In hydrogen, the effect began at 13 V. Hughes regarded the " clean-up " as due to the formation and adsorption of atoms; the rates were low, the collision efficiencies being about 0·06 at 25 V and 0·1 at 50 V in hydrogen, and 0·02 at 25 V and 0·05 at 25 V in nitrogen.

* ' Phil. Mag.,' vol. 41, p. 778 (1921).

The third possibility is that there is an actual synthesis of ammonia. In order to test this point, experiments were carried out with no applied voltage with the mixture and the pure gases. The tests used for hydrogen atoms were (1) the pressure went back to its initial value on removing the liquid air; (2) there was no decrease beyond the cooling effect on replacing the liquid air. The tests for ammonia were (1) Nessler; (2) the production of a permanent decrease approximately half the observed decrease, due to the halving of the number of molecules in the reaction $3H_2 + N_2 = 2NH_3$; (3) the ammonia could be frozen again on replacing the liquid air.

Experiments with no Applied Voltage.

Nitrogen.—The rate of fall of pressure under the conditions stated was small enough to be negligible even up to 50 V.

Hydrogen.—A large decrease in pressure was observed which did not increase very markedly on applying potentials up to 50 V. As the effect might possibly be due to electrons leaving the filament with a high velocity, retarding potentials up to 15 V were applied. Two of the pressure-time curves are shown in fig. 2. All the tests for hydrogen atoms gave positive results.

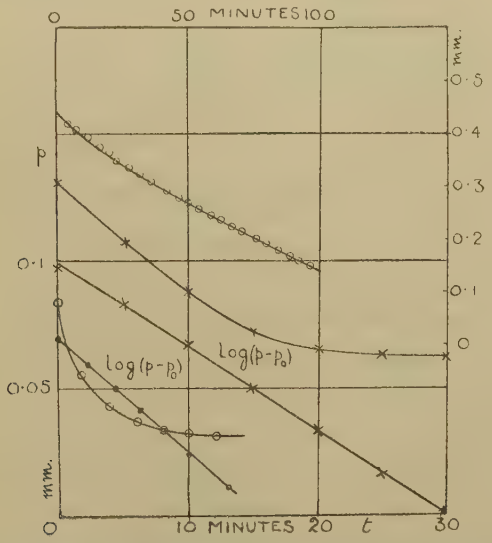


Fig. 2.—Pressure decreases with retarding potentials. The highest curve (co-ordinates top, right) is for a $N_2 + 3H_2$ mixture. The others (co-ordinates bottom, left) are for hydrogen: (*a*) circles, $p_0 = 0{\cdot}031$ mm.; (*b*) crosses, $p_0 = 0{\cdot}060$ mm.

$3H_2 + N_2$ *Mixtures.*

Continuous decreases of pressure were obtained even when retarding potentials were applied. A specimen run is shown in fig. 2. All the tests for ammonia were positive.

We may observe that the small effect with nitrogen is in agreement with Storch and Olson and Hughes, whilst the production of hydrogen atoms can only be due to dissociation either at the filament or at the electrodes or both. Evidence is given below that dissociation at the electrodes, if it occurs at all, is not the primary effect. As regards dissociation at the filament, Langmuir* has shown that hydrogen molecules are dissociated by tungsten at temperatures above about 1,300°. Rough measurements showed that the temperatures used were higher than this. On the other hand it was impossible to get workable emissions below 17 V without exceeding this temperature. Regarding the reaction without an arc as due to hydrogen atoms, the divergence between the results of Heidemann and later workers is thus explained. Starting with an arc at the lowest temperature which gives a suitable reaction rate and decreasing the potential to below 17 V, the emission becomes very small and the reaction goes very slowly. On the other hand, if the temperature is raised to keep the emission constant, hydrogen atoms are produced in abundance and a new mechanism is called forth.

Reference to fig. 2 will show that the decrease in pressure due to the hydrogen atoms decreases exponentially to zero at a pressure p_0. Plotting log $(p-p_0)$ against t gave good straight lines as shown in the graph. Hence, approximately $\log (p - p_0) = - kt + \text{const.}$ whence by differentiation $-\frac{dp}{dt} = k(p - p_0)$, showing that the rate of decrease of pressure is proportional to the difference between the pressure and the final pressure. This difference is itself proportional to the area of surface not covered by hydrogen atoms, so that it is permissible to assume that the controlling factor in this case is the surface available for accommodating hydrogen atoms.

To test this point further, between two runs the gas was merely pumped away without baking out or removing liquid air. The result is shown in the second curve, fig. 3. The initial pressure was the same, but p_0 was found to be about double, thus indicating a diminution in the area available for accommodation of hydrogen atoms due to incomplete removal of those formed in the previous run.

* 'J. Am. Chem. Soc.,' vol. 36, p. 1,708 (1914); vol. 37, p. 417 (1915); vol. 38, p. 1,145 (1916), etc.

The reactions in the hydrogen nitrogen mixture show no such saturation pressure; in fact, the reaction went on until nearly all the mixture was used up. With mixtures containing more nitrogen, the nitrogen in excess of that required for the ammonia synthesis always remained. The curves are only slightly convex to the time-axis, indicating that saturation is prevented by reaction with nitrogen to form ammonia. There was formerly a conflict of opinion as to whether normal hydrogen atoms can react with normal nitrogen molecules. Bonhoeffer* and Marshall and Taylor† found no reaction, but Hirst‡ detected N_2H_4 and NH_3 in presence of a mercury surface. Further Willey and Rideal§ found that active hydrogen produced by Wood's method (generally supposed to consist of atoms) from a discharge tube combined with N_2.

In view of this divergence between well-established experimental data, the next step was to determine the effect of the nickel or platinum electrodes on the reaction.

Experiments without an Anode.

A tube was made up containing only a tungsten filament and molybdenum wire grid. Up to 12 V only hydrogen atoms were frozen out. All the tests for hydrogen atoms were positive, tests for ammonia gave negative results. At 14 V reaction undoubtedly occurred as evidenced by the dropping of pressure until all the hydrogen and the requisite quantity of nitrogen were used up, and the other tests, including Nessler's reagent, mentioned above. A rise in the rate of reaction was also observed at 17 V, but extended results could not be obtained with this tube since there was no anode to collect the electrons, which tend to charge up the glass walls, and the emission could only be controlled by the grid leak. A composite run is shown in fig. 3 and for comparison a run with pure hydrogen under as nearly as possible identical conditions.

Experiments with an Oxide Emitter.

There was still the possibility that the 13 V point might be a critical potential of H_2 and not of H as the foregoing results tend to show. Since these experiments were carried out it was stated in a paper by Glockler, Baxter and Dalton,|| that H_2 molecules, activated by 13 V electrons, reduced copper oxide, but there is

* 'Z. Elektrochem.,' vol. 31, p. 521 (1925).

† 'J. Phys. Chem.,' vol. 23, p. 1,140 (1925).

‡ 'Proc. Camb. Phil. Soc.,' vol. 23, p. 162 (1926). See also 'Nature,' vol. 117, p. 626 (1926).

§ 'J. Chem. Soc.,' p. 669 (March, 1927).

|| 'J. Am. Chem. Soc.,' p. 58 (January, 1927).

little doubt that this is not one of the possible mechanisms of the ammonia synthesis, as the following experiments show.

An oxide emitter was made by depositing a bead of mixed SrO and BaO on a small loop of platinum. It was hoped that this would stand up in hydrogen and nitrogen since Hughes used an oxide emitter in his work on atom production. The results obtained are shown for convenience alongside the final results in fig. 5. No trace of ammonia was detected below 17 V and the rates of "clean-up" below 17 V were very small. Reaction commenced at 17 V and rose abruptly at 23 V and again at 27-28 V. At 28 V, however, the presence of water was detected by the McLeod gauge, and a blank experiment with hydrogen gave a very small pressure reduction until 28 V was reached, when a reaction rate of about the same magnitude as that in the mixed gas was observed. This point was also associated with violent arcing, so that it is evident that H^+ ions, known to be produced at this voltage, reduce BaO and SrO.

Attempts to effect complete reduction of the oxides to the metals and to use these as emitters were unsuccessful. It was therefore decided to use tungsten for examining the reaction as a whole. The oxide emitter had its use, however, since it demonstrated that the hot wire reaction in the presence of the metallic anodes and the 13 V reaction do not proceed without hydrogen atoms.

These results demonstrate conclusively that combination to form ammonia takes place between H atoms and N_2 molecules both in the presence of the nickel and platinum anodes and in the presence of 13 V electrons.

Further Experiments with a Tungsten Emitter.

Since there are two processes above 17 V, one beginning with molecular ions and the other with H atoms, in order to obtain repeatable results it was necessary to consider the H atom concentration, depending on the temperature, as well as the emission. It might be thought that the two effects could be separated; for example at 13 V for different temperatures the electronic reaction was found to be proportional to the surface action for a given emission (since both are proportional to the hydrogen atom concentration); at 17 V one would therefore expect the reaction to be the sum of a constant depending on emission alone and a reaction proportional to the surface action, depending on hydrogen atoms. The two processes are not, however, additive since they are bound to interfere. The obvious method of attack was to make each process preponderate in turn.

It was found that more nearly consistent rates at a given voltage could be obtained if the surface reaction was kept the same. The explanation of the

decrease in rate above 25 V mentioned on p. 688 was now clear. As the potential was raised the temperature had to be lowered to keep the emission constant; the reaction due to atoms was thus lessened and with it the surface reaction.

To obtain increased accuracy, it was also necessary to follow the course of the reaction at constant voltage, introducing the same mixture for the next voltage and so on. When a filament had to be replaced, a thorough baking-out was necessary until the filament had thinned to such an extent as to give the same reaction for one of the voltages already examined. Otherwise, between runs, a short baking was sufficient. In order to minimise the atom reaction, it was necessary to obtain workable emissions at low temperatures. To make the reaction with hydrogen atoms preponderate over the reaction between molecular ions, it was necessary to use higher temperatures and small emissions.

For both classes of experiment, a tube was used with a tungsten filament of length 20 mm., in the axis of a cylindrical nickel anode of diameter 48 mm., length 50 mm., and a spiral grid of molybdenum wire of diameter 0·2 mm. The difference was that for the former series, the spiral diameter was made fairly large (12 mm.), and only a few turns were used; for the latter series, the grid was placed as close as possible, and was tightly wound. Leads (of special sealing wire) were glass covered and insulation was ensured by glass spacers. The object of the grid is, of course, to provide a long path in which the electrons can collide after gaining their energy. The ideal grid would let through all the electrons accelerated to the velocity corresponding to the grid potential without any having suffered collision. Our two experimental grids can be specified as (*a*) letting through most of the electrons, but many collisions take place before the total energy is acquired. Thus definiteness is lost, but there is still a great improvement over the diode. The total emission is in this case the controlling factor. (*b*) The second grid captures a large fraction of the electrons, but most of the collisions occur outside it, hence the grid-anode current is effective, and is not too high at high temperatures.

With grid (*a*) the total emission kept constant with very little change in temperature over a wide pressure range. The total reaction and surface reaction rates were also constant.

With grid (*b*) the temperature had to be lowered to keep the emission constant as the pressure decreased. Thus, both total and surface rates decreased for a constant emission and corresponding values could be obtained.

In fig. 3 are shown a few specimen runs of the type (*a*). The method adopted was to take readings with the voltage on, and then several more at the same temperature with voltage off, and then continue with voltage on. The surface

action, which may be given the symbol S (*i.e.*, the reaction with the voltage off) was, of course, larger for the lower voltages, as the emission was kept

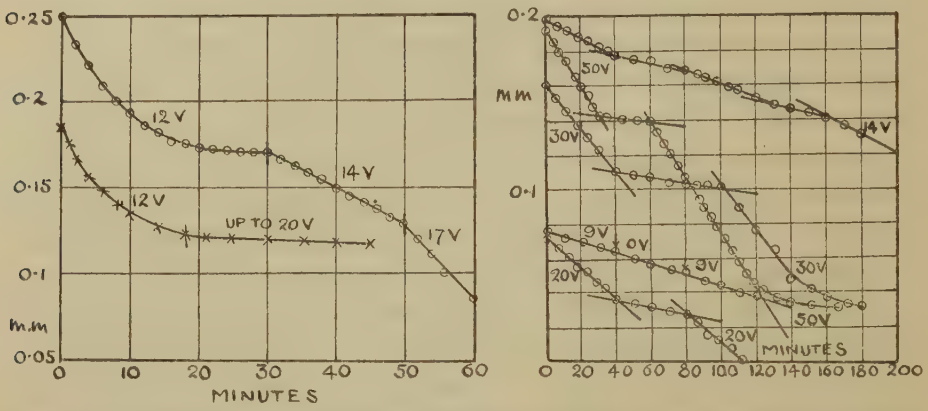


Fig. 3.—On the left, pressure decreases without an anode. Crosses represent hydrogen alone. In both sets of curves, circles represent a mixture of H_2 70 per cent. and N_2 30 per cent. On the right, total emission = 0·116 milliamp. The potential was taken off in the middle of each run.

constant throughout. The difference between the total reaction and surface reaction rates, *i.e.*, that due to excited hydrogen atoms in the bulk phase, which we may designate as H, and that due to nitrogen ions, which we may term N, the total bulk velocity being thus H + N is plotted against voltage in fig. 5 (*b*). Owing to the greater importance of the atom reaction at the lower voltages (due to the somewhat higher temperatures unavoidably used), the curve does not represent the rate of reaction as a function of voltage quantitatively, but it suffices to give the breaks.

The rates plotted in fig. 4 were found by method (*b*). Readings were taken as quickly as possible, alternately with voltage on and voltage off. On again applying the potential the temperature had to be lowered to keep the grid-anode current the same, and consequently both rates decreased. It will be seen from the figure that there was a definite total reaction for no surface reaction at the higher voltages. This represents the reaction without atoms, *i.e.*, due to molecular ions, but the accuracy was not great enough to use these intercepts as values of the true ionic bulk reaction. The gradients, less the gradient unity, for surface reaction alone are plotted on the upper curve, fig. 5 (*b*). For a few voltages, curves such as those in fig. 4 (*a*) were obtained for four or five values of the plate current. The gradients are plotted against the emission in fig. 4 (*b*). Over the range examined, the ratio of total reaction rate to surface reaction rate increased uniformly with the emission, as would

be expected on the view that the surface reaction depends on the number of hydrogen atoms, and the larger part of the total reaction apart from surface depends on the number of atoms activated by electrons.

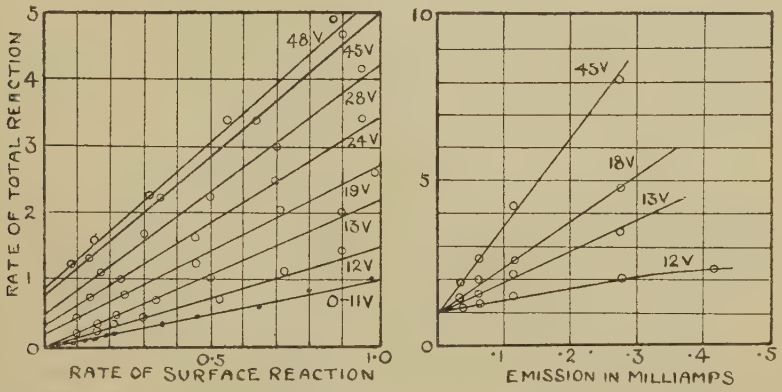


FIG. 4.—The rates are in thousandths mm. per minute. (*a*) Shows relation between reaction rate with voltage on (S + H + N) and reaction rate with voltage off (S). Partial emission = 0·116 milliamp.; (*b*) shows the proportionality of H/S to emission the gradients of curves, *i.e.* (H + S)/S, such as those in (*a*), being plotted for different values of the emission.

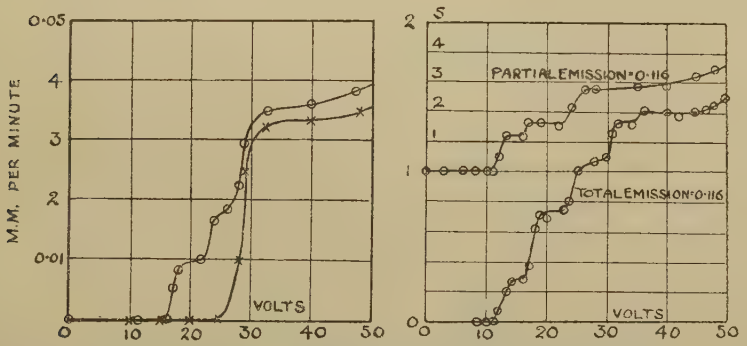


FIG. 5.—On the left are rates of decrease of pressure in mm. minute with an oxide emitter. Emission = 3·16 milliamps. Circles represent a $3H_2 + N_2$ mixture, crosses represent hydrogen alone. The lower curve on the right is obtained from the gradients of curves similar to those in fig. 3. The unit of rate of pressure fall is one-thousandth mm./minute. The upper curve represents H/S against potential.

The Efficiency of Collisions.

The number of ammonia molecules frozen out divided by the number of electrons passed should give an indication of the electron efficiency. In dealing with the rates shown in figs. 3 and 5, the total emission is taken for reasons stated

above; but since many electrons are captured by the grid without parting with their energy this will give a lower limit to the electron efficiency.

If each electron causes the freezing out of n ammonia molecules a current of C milliamps. means the removal of $\frac{nC}{e} \times 60 \times 10^{-4}$ molecules per minute. The decrease of pressure is that due to twice this number of molecules since four hydrogen and nitrogen molecules produce two molecules of ammonia. If V is the volume of gas under experimental conditions corrected to 0° C. and N the Avogadro number, N molecules produce a pressure of $\frac{22400}{V} \times 760$ mm. Hence, $120 \times 10^{-4} \frac{nC}{e}$ produce a pressure $1{\cdot}2 \times 2{\cdot}24 \times 7{\cdot}6 \times 10^4 \frac{nC}{eVN}$. Putting in $e = 1{\cdot}591 \times 10^{-20}$ and $N = 6{\cdot}062 \times 10^{23}$ and $V = 575$ c.c. (found by admitting a known volume of gas at known temperature and pressure to the reaction system under experimental conditions of temperature), we obtain a rate of decrease of pressure $= 3{\cdot}68 \times 10^{-2}\, nC$ mm./minute. For a current of 0·116 ma., the rate is equal to $4{\cdot}27 \times 10^{-3}\, n$ mm./minute. From fig. 5, at 25 V, we find a rate $= 1{\cdot}0 \times 10^{-3}$ mm./minute, so that $n = 0{\cdot}235$.

Dushman* gives the following values for the mean free paths of hydrogen and nitrogen molecules.

$$\left.\begin{matrix} H_2 & 192 \text{ mm.} \\ N_2 & 100 \text{ mm.} \end{matrix}\right\} \text{at } 0{\cdot}75 \times 10^{-3} \text{ mm. and } 25° \text{ C.}$$

The radius of the anode is 24 mm. Since the mean free path is inversely proportional to pressure, assuming the electronic mean free path to be $4\sqrt{2}$ times that of the gas (Maxwell), the electron will just collide if the pressure is

$$\frac{4\sqrt{2} \times 100}{24} \times 0{\cdot}75 \times 10^{-3}$$

or 0·018 mm. in nitrogen and about 0·034 mm. in hydrogen. The total pressure at which the straight lines in fig. 3 were obtained varied from 0·2 mm. to 0·05 mm. It would, therefore, be thought that each electron suffered at least one collision.

It was shown, however, by Compton and Van Voorhis† that the efficiency of electron collisions to produce positive molecular ions is small, being about 0·10 in H_2 and 0·055 in N_2 at 25 V. Putting in these factors brings the hydrogen pressure up to 0·34 mm. and the nitrogen pressure up to 0·33 mm.

* "High Vacuum" (1922).

† 'Phys. Rev.,' vol. 26, p. 436 (1925).

These pressures were too high to use in practice. There is, however, a general agreement between the reaction rate and rate of production of N_2^+ and H_2^+.

As regards the atom reaction, much larger rates were observed. Fig. 4 shows rates of 5×10^{-3} mm. per minute for 0·116 ma. This indicates that H atoms are much more easily activated than the molecules, and points to the H excited atom as being the chief factor in the reaction.

The Mechanism of the Reaction.

As regards the reaction at no voltage, it has been shown to take place only in the presence of hydrogen atoms and a metal surface. The reaction is thus in all probability a catalytic action at the metal surface between adsorbed hydrogen atoms and nitrogen molecules. Now, the critical radiation and ionisation potentials of hydrogen and nitrogen adsorbed on Fe, Ni, Cu and Pt surfaces have been measured by Gauger,* Wolfenden,† and Kistiakowsky.‡ The critical potentials found, 11·4 and 13·4 V for hydrogen and 11 V for nitrogen, have been attributed by them and by Taylor§ to hydrogen and nitrogen atoms produced by dissociation on active patches of the metal catalyst. Gauger, however, used a tungsten filament, and although later workers used oxide-coated platinum, the possibility of hydrogen atoms being furnished in sufficient quantity by dissociation at the hot filament was not entirely precluded. Granting that dissociation may occur to a slight extent at the cold metal surfaces, we have satisfied ourselves that nitrogen with the high heat of dissociation of not less than 150,000 calories could not be dissociated at nearly a sufficient rate to produce the high rates of reaction observed.

As regards the additional reaction beginning at 13 V, it has been shown to depend on the presence of hydrogen atoms, and can proceed in the bulk. It is, therefore, between H′ and N_2 or H and N_2'. The evidence is greatly in favour of H′. First, the hydrogen atom is known to have critical potentials at about 13 V. Again, the high collision efficiencies mentioned in the section above show that the activation is not likely to be that of N_2, particularly as atoms are known to be activated by electrons much more readily than molecules. It may be noted here that electrons do not readily dissociate molecular hydrogen although the heat of dissociation corresponds only to about 4 V. The direct evidence in

* 'J. Am. Chem. Soc.,' vol. 46, p. 674 (1924).
† 'Roy. Soc. Proc.,' vol. 110, p. 464 (1926).
‡ 'J. Phys. Chem.,' vol. 30, p. 1356 (1926).
§ 'Roy. Soc. Proc.,' vol. 113, p. 77 (1926). 'J. Phys. Chem.,' vol. 30, p. 145 (1926). 'J. Am. Chem. Soc.,' vol. 48, p. 2,840 (1926).

favour of H′ rather than N_2' is contained in the results of Hirst and Willey and Rideal, above mentioned. In the former case, there were present hydrogen atoms (by the Cario-Franck method), nitrogen and a mercury surface. In the second case "active hydrogen," formed by Wood's method was mixed with ordinary nitrogen. The "active hydrogen" from the arc would, undoubtedly, contain H′ owing to the high potential employed. It is only necessary to postulate that the life of the H′ species is sufficiently large for a small amount to reach the mixing tube from the discharge to account for the production of ammonia by a reaction between H′ and N_2.

Turning to the reaction as a whole the lower curve in fig. 5 shows breaks at 13 V, 17 V, 23 V, 30 V and a slight break at 34 V. Under the same conditions the ionisation curve for nitrogen showed breaks at about 17 V and 23 V. There is a general resemblance (except for the 13 V point) to Storch and Olson's curve. The probable explanation is that the following species are chemically reactive :—

13 V	H′
17 V	N_2^+
23 V	N^+

The other breaks are probably due to combinations of the above, *i.e.*,

$$30\ V = 13 + 17$$
$$34\ V = 2 \times 17$$

The absence of higher breaks is explained by the fact that the pressures were too low for many double efficient collisions. The upper curve of fig. 5 represents the ratio of bulk to surface action. The total rate, apart from a small ionic reaction for no surface action, *i.e.*, when the filament produced no hydrogen atoms, is proportional to the rate of surface action and to the emission (fig. 4). It, therefore, depends on the number of H′ atoms. It was surprising to find a small break at 16 V. The other (23 V–25 V) is probably a combination of a rise at 23 V (due to N^+) and at 25 V (12 + 13, since the original jumps are at 12 V and 16 V instead of 13 V and 17 V—the corrections for initial velocity, etc., are larger in these experiments owing to the higher temperatures). The absence of higher breaks is again explained by the choice of pressure range.

Some indication as to the mechanism by which N_2^+ at 17 V and N^+ at 23 V form ammonia can be obtained from the curves fig. 4. We note that the total reaction may be considered to be composed of the following, a surface reaction of velocity S, an excited hydrogen atom bulk reaction of velocity H and a third reaction due to N_2^+ at 17 volts of velocity N. The total velocity

is thus S + H + N plotted against the surface velocity S. The bulk reactions have been shown to be independent of the pressure over a wide range ; hence if the N_2^+ reacted directly with hydrogen molecules to form ammonia the curve for 19 V should be parallel to the 13 V curve cutting off an intercept on the axis. If the N_2^+ reacts with hydrogen molecules to form hydrogen atoms then both the surface and bulk reactions should increase proportionately, and the slope of the 19 V curve should again be the same as the 13 V curve, but both should pass through the same point on the axis. It will be noted, however, that neither of these conditions is fulfilled since the slope of the line at 19 volts is distinctly greater than that at 13 volts, indicating an increase in the H concentration. This may be achieved in two ways, either from hydrogen molecules by the reaction—

$$N_2^+ + e \longrightarrow N_2'$$
$$N_2' + H_2 \longrightarrow N_2 + H + H'$$

a reaction which can be shown to take place by ammonia synthesis with an oxide emitter at low temperatures ; or from hydrogen atoms by the reaction—

$$N_2^+ + e \longrightarrow N_2'$$
$$N_2 + H \longrightarrow H' + N_2 \text{ or } N_2H.$$

The main reaction between H′ and N_2 probably proceeds in the following steps—

$$N_2 + H' = N_2H$$
$$N_2H + H_2 = N_2H_2 + H$$
$$N_2H_2 + H_2 = N_2H_4$$
$$N_2H_4 + H_2 = 2NH_3$$

The observation by Storch and Olson that thinner filaments gave an increased yield of ammonia probably shows that hydrogen atoms were equally important in their experiments with low-voltage arcs. It is true that they showed one result in which a thinner filament at slightly lower temperature gave a greater rate for the same emission and potential. But they were only able to measure the *average* temperature by means of heating current and voltage drop across the filament ; and since the filament had been allowed to thin, it may be suggested that a small part of the filament was thinner and hotter than the rest (a well-known property of tungsten) and was thus producing more hydrogen atoms.

Summary.

The various possible mechanisms for the production of ammonia in a nitrogen hydrogen mixture by means of thermions have been investigated in detail. It is shown that synthesis can occur due to the following reactions—

$N_2 + H$ at the surface of platinum or nickel.

$N_2 + H'$ in the bulk at 13 volts.

The following molecular species are shown to be chemically reactive—

N_2^+ in the bulk at 17 volts,

N^+ in the bulk at 23 volts,

and possible modes of mechanism involving N_2' and H' are elaborated.

Our thanks are due to Prof. T. M. Lowry, F.R.S., who communicated this paper, and to Messrs. Brunner Mond and Co., for providing a grant to defray part of the cost of the apparatus employed.

A Contribution to the Mathematical Theory of Epidemics.

By W. O. Kermack and A. G. McKendrick.

(Communicated by Sir Gilbert Walker, F.R.S.—Received May 13, 1927.)

(From the Laboratory of the Royal College of Physicians, Edinburgh.)

Introduction.

(1) One of the most striking features in the study of epidemics is the difficulty of finding a causal factor which appears to be adequate to account for the magnitude of the frequent epidemics of disease which visit almost every population. It was with a view to obtaining more insight regarding the effects of the various factors which govern the spread of contagious epidemics that the present investigation was undertaken. Reference may here be made to the work of Ross and Hudson (1915–17) in which the same problem is attacked. The problem is here carried to a further stage, and it is considered from a point of view which is in one sense more general. The problem may be summarised as follows: One (or more) infected person is introduced into a community of individuals, more or less susceptible to the disease in question. The disease spreads from

the affected to the unaffected by contact infection. Each infected person runs through the course of his sickness, and finally is removed from the number of those who are sick, by recovery or by death. The chances of recovery or death vary from day to day during the course of his illness. The chances that the affected may convey infection to the unaffected are likewise dependent upon the stage of the sickness. As the epidemic spreads, the number of unaffected members of the community becomes reduced. Since the course of an epidemic is short compared with the life of an individual, the population may be considered as remaining constant, except in as far as it is modified by deaths due to the epidemic disease itself. In the course of time the epidemic may come to an end. One of the most important probems in epidemiology is to ascertain whether this termination occurs only when no susceptible individuals are left, or whether the interplay of the various factors of infectivity, recovery and mortality, may result in termination, whilst many susceptible individuals are still present in the unaffected population.

It is difficult to treat this problem in its most general aspect. In the present communication discussion will be limited to the case in which all members of the community are initially equally susceptible to the disease, and it will be further assumed that complete immunity is conferred by a single infection.

It will be shown in the sequel that with these reservations, the course of an epidemic is not necessarily terminated by the exhaustion of the susceptible members of the community. It will appear that for each particular set of infectivity, recovery and death rates, there exists a critical or threshold density of population. If the actual population density be equal to (or below) this threshold value the introduction of one (or more) infected person does not give give rise to an epidemic, whereas if the population be only slightly more dense a small epidemic occurs. It will appear also that the size of the epidemic increases rapidly as the threshold density is exceeded, and in such a manner that the greater the population density at the beginning of the epidemic, the smaller will it be at the end of the epidemic. In such a case the epidemic continues to increase so long as the density of the unaffected population is greater than the threshold density, but when this critical point is approximately reached the epidemic begins to wane, and ultimately to die out. This point may be reached when only a small proportion of the susceptible members of the community have been affected.

Two of the reasons commonly put forward as accounting for the termination of an epidemic, are (1) that the susceptible individuals have all been removed, and (2) that during the course of the epidemic the virulence of the causative

organism has gradually decreased. It would appear from the above results that neither of these inferences can be drawn, but that the termination of an epidemic may result from a particular relation between the population density, and the infectivity, recovery, and death rates.

Further, if one considers two populations identical in respect of their densities, their recovery and death rates, but differing in respect of their infectivity rates, it will appear that epidemics in the population with the higher infectivity rate may be great as compared with those in the population with the lower infectivity rate, especially if the density of the former population is in the neighbourhood of the threshold value. If, then, the density of a particular population is normally very close to its threshold density it will be comparatively free from epidemic, but if this state is upset, either by a slight increase in population density, or by a slight increase in the infectivity rate, a large epidemic may break out. Such great sensitiveness of the magnitude of the epidemic with respect to these two factors, may help to account for the apparently sporadic occurrence of large epidemics, from very little apparent cause. Further, it will appear that a similar state of affairs holds with respect to diseases which are transmitted through an intermediate host. In this case the product of the two population densities is the determining factor, and no epidemic can occur when the product falls below a certain threshold value.

General Theory.

(2) We shall first consider the equations which arise when the time is divided into a number of separate intervals, and infections are supposed to take place only at the instant of passing from one interval to the next, and not during the interval itself. We shall take the size of this interval, which at present may be considered constant, as the unit of time, and we shall denote the number of individuals in unit area at the time t who have been infected for θ intervals by $v_{t,\theta}$. The total number who are ill at this interval t is $\sum_{\theta=0}^{t} v_{t,\theta}$, which we shall call y_t. It should be noted that $v_{t,0}$ denotes the number of individuals at the time t who are at the beginning of their infection. Also we shall use the symbol v_t to denote the number who actually undergo the process of infection during the transition from the interval $t-1$, to the interval t. In general $v_{t,0} = v_t$ except at the origin, where we assume that a certain number y_0 of the population have just been infected, although this infection is naturally dependent on some process outside that defined by the equations which we shall develop. Thus

$$v_{0,0} = v_0 + y_0. \qquad (1)$$

The whole process is indicated in the following schema :—

Fresh infections.	Numbers at each stage of illness.	Number ill.
v_3	$v_{3,0}$ $v_{3,1}$ $v_{3,2}$ $v_{3,3}$	y_3
v_2	$v_{2,0}$ $v_{2,1}$ $v_{2,2}$	y_2
v_1	$v_{1,0}$ $v_{1,1}$	y_1
v_0	$v_{0,0}$	y_0

The arrows indicate the course followed by each individual until he recovers or dies.

If ψ_θ denotes the rate of removal, that is to say it is the sum of the recovery and death rates, then the number who are removed from each θ group at the end of the interval t is $\psi_\theta v_{t,\theta}$, and this is clearly equal to $v_{t,\theta} - v_{t+1,\theta+1}$.
Thus

$$\begin{aligned} v_{t,\theta} &= v_{t-1,\theta-1}(1 - \psi(\theta - 1)) \\ &= v_{t-2,\theta-2}(1 - \psi(\theta - 1)(1 - \psi(\theta - 2)) \\ &= v_{t-\theta,0}\, B_\theta, \end{aligned} \tag{2}$$

where B_θ is the product $(1 - \psi(\theta - 1))(1 - \psi(\theta - 2)) \ldots (1 - \psi(0))$.

Now v_t denotes the number of persons in unit area who became infected at the interval t, and this must be equal to $x_t \sum_1^t \phi_\theta v_{t,\theta}$ where x_t denotes the number of individuals still unaffected, and ϕ_θ is the rate of infectivity at age θ. (It is indifferent whether we include the term $\phi_0 v_{t,0}$ or not, since in this paper we assume that ϕ_0 is zero, that is that an individual is not infective at the moment of infection.) This follows since the chance of an infection is proportional to the number of infected on the one hand, and to the number not yet infected on the other.

It is clear that

$$\begin{aligned} x_t &= N - \sum_0^t v_{t,0} \\ &= N - \sum_0^t v_t - y_0, \end{aligned} \tag{3}$$

where N is the initial population density.

If z_t denotes the number who have been removed by recovery and death, then

$$x_t + y_t + z_t = N. \tag{4}$$

Thus we have

$$v_t = x_t \sum_1^t \phi_\theta v_{t,\theta} = x_t \sum_1^t \phi_\theta B_\theta v_{t-\theta,0} \qquad \text{(by 2)}$$

$$= x_t \left(\sum_1^t A_\theta v_{t-\theta} + A_t y_0\right) \qquad \text{(by 1)}, \tag{5}$$

where A_θ is written for $\phi_\theta B_\theta$.

Also

$$y_t = \sum_0^t v_{t,\theta} = \sum_0^t B_\theta v_{t-\theta} + B_t y_0. \tag{6}$$

By definition

$$-v_t = x_{t+1} - x_t, \tag{7}$$

hence equation (5) may be written

$$x_t - x_{t+1} = x_t \left(\sum_1^t A_\theta v_{t-\theta} + A_t y_0\right). \tag{8}$$

Also $z_{t+1} - z_t$ is the number of persons who are removed at the end of the interval of time t, and this is equal to $\sum_1^t \psi_\theta v_{t,\theta}$, *i.e.*, to $\sum_1^t \psi_\theta B_\theta v_{t-\theta} + \psi_t B_t y_0$, hence writing C_θ for $\psi_\theta B_\theta$ we have

$$z_{t+1} - z_t = \sum_1^t C_\theta v_{t-\theta} + C_t y_0. \tag{9}$$

Also by (4)

$$y_{t+1} - y_t = x_t \left[\sum_1^t A_\theta v_{t-\theta} + A_t y_0\right] - \left[\sum_1^t C_\theta v_{t-\theta} + C_t y_0\right]. \tag{10}$$

(3) If now we allow the subdivisions of time to increase in number so that each interval becomes very small, then in the limit the above equations (4, 7, 8, 9) become

$$x_t + y_t + z_t = N. \tag{11}$$

$$v_t = -\frac{dx_t}{dt}, \tag{12}$$

$$\frac{dx_t}{dt} = -x_t \left[\int_0^t A_\theta v_{t-\theta}\, d\theta + A_t y_0\right], \tag{13}$$

$$\frac{dz_t}{dt} = \int_0^t C_\theta v_{t-\theta}\, d\theta + C_t y_0, \tag{14}$$

and from (6)

$$y_t = \int_0^t B_\theta v_{t-\theta}\, d\theta + B_t y_0, \tag{15}$$

where

$$B_\theta = e^{-\int_0^\theta \psi(a)\, da}, \qquad A_\theta = \phi_\theta B_\theta, \qquad \text{and} \qquad C_\theta = \psi_\theta B_\theta.$$

It can, however, be shown that these five relations are not independent and in fact that (11) is a necessary consequence of (13), (14) and (15). The four

independent relations (12), (13), (14) and (15) determine the four functions x, y, z and v.

By equation (13), dropping the suffix t except when necessary in the analysis,

$$\begin{aligned}\frac{dx}{dt} &= -x\left[\int_0^t \mathrm{A}_\theta v_{t-\theta}\,d\theta + \mathrm{A}_t y_0\right],\\ &= -x\left[\int_0^t \mathrm{A}_{t-\theta} v_\theta\,d\theta + \mathrm{A}_t y_0,\right]\\ &= x\left[\int_0^t \mathrm{A}_{t-\theta}\frac{dx_\theta}{d\theta}d\theta - \mathrm{A}_t y_0\right],\end{aligned}$$

where x in the integral is now a function of θ.

Therefore

$$\begin{aligned}\frac{d\log x}{dt} &= \mathrm{A}_{t-\theta}x_\theta\Big|_0^t - \int_0^t x_\theta\frac{d\,\mathrm{A}_{t-\theta}}{d\theta}d\theta - \mathrm{A}_t y_0,\\ &= \mathrm{A}_0 x_t - \mathrm{A}_t x_0 + \int_0^t x_\theta \mathrm{A}'_{t-\theta}\,d\theta - \mathrm{A}_t y_0,\end{aligned}$$

where

$$\mathrm{A}'_{t-\theta} = \frac{d\,\mathrm{A}_{t-\theta}}{d\,(t-\theta)} = -\frac{d\mathrm{A}_{t-\theta}}{d\theta}.$$

But $\mathrm{A}_0 = \phi_0\mathrm{B}_0 = \phi_0 = 0$, since we assume that an individual at the moment of becoming infected cannot transmit infection.

Hence

$$\left.\begin{aligned}\frac{d\log x}{dt} &= -\mathrm{A}_t(x_0+y_0) + \int_0^t x_\theta \mathrm{A}'_{t-\theta}\,d\theta,\\ &= -\mathrm{A}_t\mathrm{N} + \int_0^t \mathrm{A}'_\theta x_{t-\theta}\,d\theta.\end{aligned}\right\} \qquad (16)$$

We have not been able to solve this equation in such a way as to give x in terms of t as an explicit function. It may, however, be pointed out that this is an integral equation similar to Volterra's equation

$$f(t) = \phi(t) + \int_0^t \mathrm{N}(t,\,\theta)\,\phi(\theta)\,d\theta,$$

except that in place of $f(t)$ we have $\dfrac{d\log f(t)}{dt}$.

(4) If we consider an equation of the form

$$\frac{d\log x}{dt} = \mathrm{A}_t + \lambda\int_0^t \mathrm{N}(t,\,\theta)\,x(\theta)\,d\theta,$$

of which the above equation is a particular example, it would appear that a solution can be arrived at by a series of successive approximations in a way similar to the method used in resolving Volterra's equation.

We may write

$$x = f_0(t) + \lambda f_1(t) + \lambda^2 f_2(t) + \text{etc.}$$

It is easily seen that after substituting this expression in the equation

$$\frac{dx}{dt} = x\left[\mathrm{A}_t + \lambda \int_0^t \mathrm{N}(t, \theta)\, x(\theta)\, d\theta\right],$$

and equating the coefficients of the powers of λ, we obtain

$$\frac{d}{dt} f_n(t) = f_n(t)\,\mathrm{A}_t + f_{n-1}(t) \int_0^t \mathrm{N}(t, \theta) f_0(\theta)\, d\theta + f_{n-2}(t) \int_0^t \mathrm{N}(t, \theta) f_1(\theta)\, d\theta$$

$$+ \ldots + f_0(t) \int_0^t \mathrm{N}(t, \theta) f_{n-1}(\theta)\, d\theta$$

$$= \mathrm{L}_{n-1}(t) \quad \text{say.}$$

This is a differential equation for $f_n(t)$ of which the solution is

$$f_n(t)\, e^{-\int_0^t \mathrm{A}_t dt} = \int_0^t \mathrm{L}_{n-1}(t)\, e^{-\int_0^t \mathrm{A}_t dt}\, dt + \text{constant},$$

where $\mathrm{L}_{n-1}(t)$ is a function of the f's.

Also $f_n(0)$ is zero ($n > 0$), since the initial conditions are presumably independent of λ. Hence the constants of integration are all zero except $f_0(0)$.

In the case of this function we have

$$\frac{df_0(t)}{dt} = f_0(t)\,\mathrm{A}_t,$$

whence

$$f_0(t) = f_0(0)\, e^{\int_0^t \mathrm{A}_t dt},$$

so that $f_0(0) = x_0$.

We thus have for the solution of the integral equation,

$$x = x_0 \mathrm{E}_t + \sum_{n=1}^{\infty} \lambda^n \mathrm{E}_t \int_0^t \frac{\mathrm{L}_{n-1}(t)}{\mathrm{E}_t}\, dt,$$

$$= \mathrm{E}_t\left[x_0 + \sum_1^{\infty} \lambda^n \int_0^t \frac{\mathrm{L}_{n-1}(t)}{\mathrm{E}_t}\, dt\right],$$

where E_t is written for exp. $\int_0^t \mathrm{A}_t\, dt$; and when $\lambda = 1$

$$x = \mathrm{E}_t\left[x_0 + \sum_1^{\infty} \int_0^t \frac{\mathrm{L}_{n-1}(t)}{\mathrm{E}_t}\, dt\right]. \qquad (17)$$

(5) Returning to equation (16) let us consider it in the rather more general form

$$\frac{d \log x}{dt} = \mathrm{A}_t + \int_0^t \mathrm{Q}_{t-\theta} x_\theta d\theta.$$

Multiplying both sides by e^{-zt} where the real part of z is positive, and integrating with respect to t between the limits zero and infinity, we have

$$\int_0^\infty e^{-zt} \frac{d \log x}{dt} dt = \int_0^\infty e^{-zt} \mathrm{A}_t dt + \int_0^\infty e^{-zt} \int_0^t \mathrm{Q}_{t-\theta} x_\theta \, d\theta \, dt,$$

therefore

$$-\log x_0 + \int_0^\infty z e^{-zt} \log x \, dt = \mathrm{F}(z) + \int_0^\infty e^{-z\theta} \mathrm{Q}_\theta \, d\theta \int_0^\infty e^{-zt} x_t \, dt,$$

$$= \mathrm{F}(z) + \mathrm{F}_1(z) \int_0^\infty e^{-zt} x_t \, dt,$$

where $\mathrm{F}(z)$ is written for $\int_0^\infty e^{-zt} \mathrm{A}_t dt$, and $\mathrm{F}_1(z)$ for $\int_0^\infty e^{-z\theta} \mathrm{Q}_\theta \, d\theta$. Clearly $e^{-zt} \log x$ tends to zero as t tends to infinity, whilst x never exceeds the initial value $\mathrm{N} - y_0$.

Thus

$$\int_0^\infty e^{-zt} (z \log x - \mathrm{F}_1(z) x) \, dt = \mathrm{F}(z) + \log x_0. \tag{18}$$

It will be seen that this is an equation of the form

$$\int_0^\infty \phi(x, z) \, \psi(z, t) \, dt = \chi(z), \tag{19}$$

where the functions ϕ, ψ and χ are known, and x is a function of t. z may have any value provided that its real part is positive. It follows that the formal solution obtained in the previous paragraph, equation (17), must satisfy this equation (19). If $\phi(x, z)$ had not contained z explicitly equation (19) would be of Fredholm's first type. From this point of view the above equation may be regarded as a generalisation of Fredholm's equation of the first type.

(6) Let us now integrate equation (13) with respect to t, between the limits zero and infinity.

We have

$$-\int_0^\infty \frac{d \log x}{dt} \, dt = \int_0^\infty \int_0^t \mathrm{A}_\theta v_{t-\theta} \, d\theta \, dt + y_0 \int_0^\infty \mathrm{A}_t \, dt,$$

hence

$$\log \frac{x_0}{x_\infty} = \int_0^\infty \mathrm{A}_\theta \, d\theta \int_0^\infty v_t \, dt + y_0 \int_0^\infty \mathrm{A}_t \, dt.$$

If we put A for $\int_0^\infty A_t dt$, and use the relation $\int_0^\infty v_t dt = -\int_0^\infty \frac{dx}{dt} dt = x_0 - x_\infty$

we have

$$\log \frac{x_0}{x_\infty} = A(x_0 - x_\infty) + Ay_0 = A(N - x_\infty).$$

Let us introduce the value $p = \frac{N - x_\infty}{N}$, so that p is the proportion of the population who become infected during the epidemic.

Then $x_\infty = N(1 - p)$

and

$$-\log \frac{1-p}{1-\frac{y_0}{N}} = ANp. \tag{20}$$

This equation determines the size of the epidemic in terms of A, N, and y_0, and we shall make use of it later.

If we treat equation (15) in a similar manner, we obtain the relation

$$\int_0^\infty y_t dt = N p \int_0^\infty B_\theta d\theta.$$

Thus $\int_0^\infty B_\theta d\theta$ is the average case duration.

(7) Finally the observational data are given in terms of x, y and z, though in particular instances the information may be incomplete. The problem may arise of obtaining A_θ and B_θ as functions of θ, and thus of acquiring knowledge regarding ϕ_θ and ψ_θ, the infectivity and removal rates.

In equation (13) v_t and $d \log x/dt$ are known functions of t and so the equation is of the type discussed by Fock (1924). We shall apply his method to obtain the solution of this and similar equations.

By equation (13)

$$-\int_0^\infty e^{-zt} \frac{d \log x}{dt} dt = \int_0^\infty e^{-zt} \int_0^t A_\theta v_{t-\theta} d\theta \, dt + y_0 \int_0^\infty e^{-zt} A_t dt,$$

$$= \int_0^\infty e^{-z\theta} A_\theta d\theta \int_0^\infty e^{-zt} v_t dt + y_0 \int_0^\infty e^{-zt} A_t dt,$$

therefore

$$\int_0^\infty e^{-zt} A_t dt = \frac{-\int_0^\infty e^{-zt} \frac{d \log x}{dt} dt}{y_0 + \int_0^\infty e^{-zt} v_t dt}, \tag{21}$$

and we shall denote this last expression by the symbol $F_2(z)$ whence

$$A_\theta = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} e^{zt} F_2(z)\, dz. \tag{21A}$$

By equation (15)

$$\int_0^\infty e^{-zt} y_t\, dt = \int_0^\infty e^{-zt}\int_0^t B_\theta v_{t-\theta}\, d\theta\, dt + y_0 \int_0^\infty e^{-zt} B_t\, dt,$$

whence

$$\int_0^\infty e^{-zt} B_t\, dt = \frac{\int_0^\infty e^{-zt} y_t\, dt}{y_0 + \int_0^\infty e^{-zt} v_t\, dt}, \tag{22}$$

we shall denote this last expression by $F_3(z)$, and so

$$B_\theta = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} e^{zt} F_3(z)\, dt. \tag{22A}$$

Equations (21A) and (22A) give A_θ and B_θ in terms of the observable data.

If $F_2(z)$ and $F_3(z)$ can be expressed as rational functions of z, then in place of Laplace's transformation we can use the simpler solution given in the next section.

Special Cases.

A.—*The earlier stages of an epidemic in a large population.*

(8) During the early stages of an epidemic in a large population, the number of unaffected persons may be considered to be constant, since any alteration is small in comparison with the total number. Equation (13) becomes

$$-\frac{dx}{dt} = v_t = N\left[\int_0^\infty A_\theta v_{t-\theta}\, d\theta + A_t y_0\right],$$

where N is this constant population per unit area.

Using Fock's method

$$\int_0^\infty e^{-zt} v_t\, dt = \frac{N y_0 \int_0^\infty e^{-zt} A_t\, dt}{1 - N\int_0^\infty e^{-zt} A_t\, dt} \tag{23}$$

and we shall denote this by $F_4(z)$.

Thus

$$v_t = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} e^{zt} F_4(z)\, dz. \tag{23A}$$

Making use of equation (15) we have similarly

$$\int_0^\infty e^{-zt} y_t\,dt = \int_0^\infty e^{-zt} \int_0^t \mathrm{B}_\theta v_{t-\theta}\,d\theta\,dt + y_0 \int_0^\infty e^{-zt}\mathrm{B}_t\,dt,$$

$$= \int_0^\infty e^{-zt} v_t\,dt \int_0^\infty e^{-z\theta}\mathrm{B}_\theta\,d\theta + y_0 \int_0^\infty e^{-zt}\mathrm{B}_t\,dt,$$

$$= \frac{\mathrm{N}y_0 \int_0^\infty e^{-zt}\mathrm{A}_t\,dt \int_0^\infty e^{-zt}\mathrm{B}_t\,dt}{1 - \mathrm{N}\int_0^\infty e^{-zt}\mathrm{A}_t\,dt} + y_0 \int_0^\infty e^{-zt}\mathrm{B}_t\,dt,$$

$$= \frac{y_0 \int_0^\infty e^{-zt}\mathrm{B}_t\,dt}{1 - \mathrm{N}\int_0^\infty e^{-zt}\mathrm{A}_t\,dt} \qquad (24)$$

which we shall call $\mathrm{F}_5(z)$.

Thus

$$y_t = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} e^{zt}\mathrm{F}_5(z)\,dz. \qquad (24\mathrm{A})$$

Further we may find the integral equation for y_t as follows :—

$$y_t = \int_0^t \mathrm{B}_{t-\theta} v_\theta\,d\theta + \mathrm{B}_t y_0,$$

$$= \mathrm{N}\int_0^t \mathrm{B}_{t-\theta}\left(\int_0^\theta \mathrm{A}_{\theta-z} v_z\,dz + \mathrm{A}_\theta y_0\right) d\theta + \mathrm{B}_t y_0,$$

$$= \mathrm{N}\int_0^t \mathrm{B}_{t-\theta}\int_0^\theta \mathrm{A}_{\theta-z} v_z\,dz\,d\theta + \mathrm{N}y_0\int_0^t \mathrm{B}_{t-\theta}\mathrm{A}_\theta\,d\theta + \mathrm{B}_t y_0,$$

$$= \mathrm{N}\int_0^t \mathrm{A}_{t-\theta}\int_0^\theta \mathrm{B}_{\theta-z} v_z\,dz\,d\theta + \mathrm{N}y_0\int_0^t \mathrm{A}_{t-\theta}\mathrm{B}_\theta\,d\theta + \mathrm{B}_t y_0,$$

$$= \mathrm{N}\int_0^t \mathrm{A}_{t-\theta}\,(y_\theta - \mathrm{B}_\theta y_0 + \mathrm{B}_\theta y_0)\,d\theta + \mathrm{B}_t y_0,$$

$$= \mathrm{N}\int_0^t \mathrm{A}_{t-\theta} y_\theta\,d\theta + \mathrm{B}_t y_0. \qquad (25)$$

It is easy to show that by solving this directly we obtain the solution (24).

In a previous communication, McKendrick (1925–26), these solutions were given in a somewhat different form. The equation for $v_{t,0}$ was given as

$$v_{t,0} = \int_0^t \mathrm{A}_\theta v_{t-\theta,0}\,d\theta,$$

and the solution obtained was

$$v_{t,0} = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} e^{zt}\,\frac{N_0}{1-\int_0^\infty e^{-z\theta}A_\theta\,d\theta}\,.\,dz.$$

It was remarked that $v_{t,0}$ had a singularity at the point $t=0$. In the present discussion we regard the original infections as occurring at the very beginning of the epidemic but in such a way as to be independent of the equations which define the epidemic proper. Thus $v_{t,0} = v_t$ except in the short interval of time 0 to ϵ, and during this interval the integral equation does not hold, but instead $\int_0^\epsilon v_{t,0}\,dt$ is equal to y_0.

Thus

$$\begin{aligned} v_{t,0} &= v_{t,0} - v_{\epsilon,0} + v_{\epsilon,0}, \\ &= \int_\epsilon^t A_{t-\theta}v_{\theta,0}\,d\theta + \int_0^\epsilon A_{t-\theta}v_{\theta,0}\,d\theta, \\ &= \int_0^t A_{t-\theta}v_\theta\,d\theta + A_{t-\epsilon'}\int_0^\epsilon v_{\theta,0}\,d\theta, \quad \text{where} \quad 0<\epsilon'<\epsilon, \\ &= \int_0^t A_{t-\theta}v_\theta\,d\theta + A_t y_0. \end{aligned}$$

Thus the integral equation previously given for $v_{t,0}$ implies the equation now given for v_t. The solution previously given may be written in the form

$$v_{t,0} = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} e^{zt}F(z)\,dz,$$

where

$$F(z) = \frac{y_0}{1-\int_0^\infty e^{-z\theta}A_\theta\,d\theta} : \text{ let us denote this by } \frac{y_0}{1-A}.$$

In the new form

$$F_A(z) = -y_0 + \frac{y_0}{1-A} = \frac{Ay_0}{1-A},$$

which is the same as in equation (23) when one notes that in the former discussion the function A was taken as including N. Now if v_t has no singularities, the Laplacian solution of $F_A(z)$ is a function with no singularities and so the Laplacian of y_0 corresponds to the singularity. It is easy to see that the Laplacian solution $\frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} e^{zt}(-y_0)\,dz$ corresponds to a function $\phi(t)$ such that $\int_0^\infty e^{-zt}\phi(t)\,dt = -y_0$. Now if $\phi(t)$ is zero from ϵ to ∞, and becomes infinite at the origin in such a way that $\int_0^\epsilon \phi(t)\,dt$ tends to y_0 as ϵ tends to zero, then it is clear that the above equation will be true. And so the expression $\int_{a-i\infty}^{a+i\infty} e^{zt}(-y_0)\,dz$ may be taken as representing a function with exactly the same properties as $v_t - v_{t,0}$. That is to say it is zero from ϵ to ∞ and $\int_0^\epsilon (v_t - v_{t,0})\,dt = -y_0$, when ϵ becomes very small.

These values of v_t and y_t constitute the general solution of the problem in the case where N is considered as remaining constant, if A_θ and B_θ, or ϕ_θ and ψ_θ are given.

We can as before readily obtain the values A_θ and B_θ from observed values of v_t and y_t, and we find

$$A_\theta = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} e^{zt}\frac{\int_0^\infty e^{-zt}v_t\,dt}{Ny_0 + N\int_0^\infty e^{-zt}v_t\,dt}\,dz, \tag{26}$$

and

$$B_\theta = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} e^{zt}\frac{\int_0^\infty e^{-zt}y_t\,dt}{y_0 + \int_0^\infty e^{-zt}v_t\,dt}\,dz. \tag{27}$$

For the arithmetical solution of the integral equations the reader is referred to Whittaker (' Roy. Soc. Proc.,' A, vol. 94, p. 367, 1918).

(9) It will be observed that solutions (21, 22, 23, 24, 26, 27) depend upon an equation of the type $\int_0^\infty e^{-zt}\phi(t)\,dt = F(z)$ whose solution can be expressed by the use of Laplace's transformation.

If $F(z)$ can be expressed as a rational function of the form $\frac{\psi_n(z)}{\psi_m(z)}$ where ψ_n and ψ_m are polynomials of degree n and m respectively, and n is less than m, then it is always possible to express $F(z)$ in the form $\Sigma\Sigma\frac{A_{r,s}}{(z-\alpha_r)^s}$ where r and s vary from unity to a and b respectively, and a and b have finite values.

But

$$\int_0^\infty e^{-zt}e^{\alpha t}t^c\,dt = \frac{c!}{(z-\alpha)^{c+1}},$$

hence a solution of

$$\int_0^\infty e^{-zt}\phi(t)\,dt = \Sigma\Sigma\frac{A_{r,s}}{(z-\alpha_r)^s}$$

is given by

$$\phi(t) = \Sigma\Sigma\frac{A_{r,s}}{(s-1)!}t^{s-1}e^{\alpha_r t}: \qquad \text{see Fock (\textit{loc. cit.}).} \tag{28}$$

B. *Constant Rates.*

(10) Much insight can be obtained as to the process by which epidemics in limited populations run their peculiar courses, and end in final extinction,

from the consideration of the special case in which ϕ and ψ are constants κ and l respectively.

In this case the equations are

$$\left.\begin{aligned} \frac{dx}{dt} &= -\kappa xy \\ \frac{dy}{dt} &= \kappa xy - ly \\ \frac{dz}{dt} &= ly \end{aligned}\right\} \tag{29}$$

and as before $x + y + z = \mathrm{N}$.

Thus

$$\frac{dz}{dt} = l(\mathrm{N} - x - z),$$

and $\frac{dx}{dz} = -\frac{\kappa}{l}x$, whence $\log \frac{x_0}{x} = \frac{\kappa}{l}z$, since we assume that z_0 is zero.

Thus

$$\frac{dz}{dt} = l\left(\mathrm{N} - x_0 e^{-\frac{\kappa}{l}z} - z\right).$$

Since it is impossible from this equation to obtain z as an explicit function of t, we may expand the exponential term in powers of $\frac{\kappa}{l}z$, and we shall assume that $\frac{\kappa}{l}z$ is small compared with unity.

Thus

$$\frac{dz}{dt} = l\left\{\mathrm{N} - x_0 + \left(\frac{\kappa}{l}x_0 - 1\right)z - \frac{x_0\kappa^2 z^2}{2l^2}\right\}.$$

But $\mathrm{N} - x_0 = y_0$, where y_0 is small. It is for this reason that we have to take into consideration the third term in z^2, as although $\frac{\kappa}{l}z$ is small compared with unity, its square may not be small as compared with $\left(\frac{\kappa}{l}x_0 - 1\right)z$.

The solution of this equation is

$$z = \frac{l^2}{\kappa^2 x_0}\left\{\frac{\kappa}{l}x_0 - 1 + \sqrt{-q}\tanh\left(\frac{\sqrt{-q}}{2}lt - \phi\right)\right\} \tag{30}$$

where

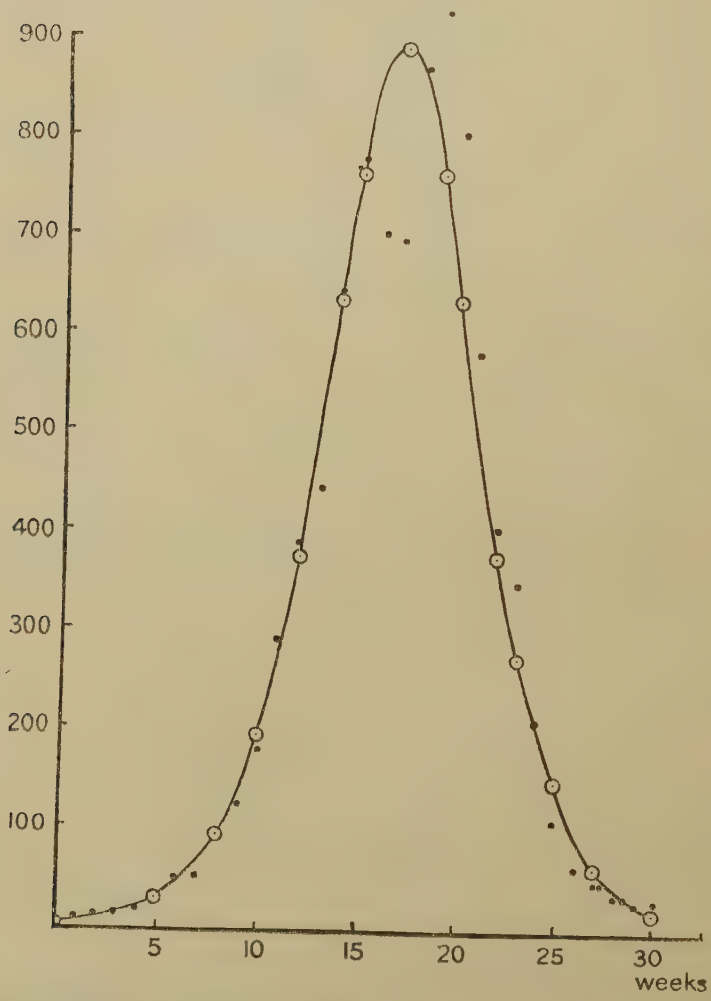
$$\phi = \tanh^{-1}\frac{\frac{\kappa}{l}x_0 - 1}{\sqrt{-q}},$$

and

$$\sqrt{-q} = \left\{\left(\frac{\kappa}{l}x_0 - 1\right)^2 + 2x_0 y_0 \frac{\kappa^2}{l^2}\right\}^{\frac{1}{2}}.$$

Also for the rate at which cases are removed by death or recovery which is the form in which many statistics are given

$$\frac{dz}{dt} = \frac{l^3}{2x_0\kappa^2}\sqrt{-q}\,\mathrm{sech}^2\left(\frac{\sqrt{-q}}{2}lt - \phi\right). \tag{31}$$



The accompanying chart is based upon figures of deaths from plague in the island of Bombay over the period December 17, 1905, to July 21, 1906. The ordinate represents the number of deaths per week, and the abscissa denotes the time in weeks. As at least 80 to 90 per cent. of the cases reported terminate fatally, the ordinate may be taken as approximately representing dz/dt as a function of t. The calculated curve is drawn from the formula

$$\frac{dz}{dt} = 890\,\mathrm{sech}^2\,(0{\cdot}2t - 3{\cdot}4).$$

We are, in fact, assuming that plague in man is a reflection of plague in rats, and that with respect to the rat (1) the uninfected population was uniformly susceptible; (2) that all susceptible rats in the island had an equal chance of being infected; (3) that the infectivity, recovery, and death rates were of constant value throughout the course of sickness of each rat; (4) that all cases ended fatally or became immune; and (5) that the flea population was so large that the condition approximated to one of contact infection. None of these assumptions are strictly fulfilled and consequently the numerical equation can only be a very rough approximation. A close fit is not to be expected, and deductions as to the actual values of the various constants should not be drawn. It may be said, however, that the calculated curve, which implies that the rates did not vary during the period of the epidemic, conforms roughly to the observed figures.

Further at the end of the epidemic

$$z = \frac{2l}{\kappa x_0}\left(x_0 - \frac{l}{\kappa}\right) \tag{32}$$

where y_0 has been neglected. This is obviously no limitation as y_0, the initial number of infected cases is usually small as compared with x_0. It is clear that when x_0, which is identical with N if y_0 be neglected, is equal to l/κ, no epidemic can take place. If, however, N slightly exceeds this value then a small epidemic will occur, and if we write $\mathrm{N} = \frac{l}{\kappa} + n$, its magnitude will be

$$2\frac{l}{\kappa}\frac{n}{\mathrm{N}} \quad \text{or} \quad 2n - \frac{2n^2}{\mathrm{N}}.$$

In this sense the population density $\mathrm{N}_0 = \frac{l}{\kappa}$ may be considered as the threshold density of the population for an epidemic with these characteristics. No epidemic can occur unless the population density exceeds this value, and if it does exceed the threshold value then the size of the epidemic will be, to a first approximation, equal to $2n$, that is to twice the excess (if n is small as compared with N). And so at the end of the epidemic the population density will be just as far below the threshold density, as initially it was above it.

At first sight it appears peculiar that in such a homogeneous population the epidemic should at first increase and then diminish. The reason for this behaviour is readily appreciated when attention is focussed on the conditions obtaining when the epidemic is at its maximum. By equation (29) this occurs when $\frac{dy}{dt} = 0$, that is when $x = \frac{l}{\kappa}$, or when the unaffected population has been reduced to its threshold value. Once the population is below this value, any particular infected individual has more chance of being removed by recovery or by death than of becoming a source of further infection, and so the epidemic

commences to decrease. In fact, as remarked above, in small epidemics the curve for y is symmetrical about the maximum. This symmetry exists for y as a function of t, and consequently also for dz/dt, that is to say the curve of removal by recovery or by death. On the other hand no such symmetry is obtained in the curve of case incidence, that is of $-\frac{dx}{dt} = \kappa xy$. This is clear since y is symmetrical and $x = e^{-\frac{1}{\kappa}\int y\,dt}$.

C. *Magnitude of small epidemics in general case.*

(11) We have seen that in the case last discussed, that is where the population is limited, and the characteristic rates are constants, a threshold value exists, such that no epidemic can arise if the density is below this value, whereas if the density be above it, the size of the epidemic is equal to twice the excess, provided that the excess be a small fraction of the threshold density. It is of importance to enquire how far a similar result is true in the general case where the characteristic rates vary during the course of the disease.

We found that

$$-\log\frac{1-p}{1-\frac{y_0}{N}} = ApN, \tag{20}$$

where p is the proportion of the population infected during the epidemic, and

$$A = \int_0^\infty A_\theta\,d\theta = \int_0^\infty \phi_\theta e^{-\int_0^\theta \psi_\alpha\,d\alpha}\,d\theta.$$

We shall assume that y_0/N is small as compared with unity, and can be neglected.

It is clear that when p is greater than zero, $-\log(1-p) > p$, hence $ApN > p$ and consequently $AN > 1$.

That is to say for an epidemic to occur (that is for p to be greater than zero), N must be greater than $1/A$. Writing $N_0 = 1/A$ and $N = N_0 + n$ we have

$$p + \frac{p^2}{2} + \frac{p^3}{3} + \ldots = ApN$$
$$= p\left(1 + \frac{n}{N_0}\right),$$

hence

$$\frac{p}{2} + \frac{p^2}{3} + \ldots = \frac{n}{N_0},$$

or neglecting powers of p higher than the first

$$pN = 2n\frac{N}{N_0} = 2n\left(1 + \frac{n}{N_0}\right) = 2n, \tag{33}$$

approximately, as n/N_0 may be neglected as compared with unity.

A difficulty occurs due to the fact that y_0 can have no value less than unity, and so y_0/N cannot be made indefinitely small. It appears, in fact, that under certain conditions quite a number of cases might occur at the threshold value, but these would be sporadic cases and would not constitute an epidemic in the true sense. The difficulty may be got over if we allow the unit of area to increase. If we increase it κ times then N_0 increases to κN_0 and A becomes A/κ, so that AN_0 does not change. On the other hand y_0/N_0 becomes $y_0/\kappa N_0$, and although y_0 can never be less than unity, κ can be made indefinitely large, and so $y_0/\kappa N_0$ may ultimately be neglected as compared with unity.

It thus appears that precisely the same result is arrived at in this case, as in the simpler case in which the rates were constants. There exists a threshold population whose density is equal to 1/A, and when an epidemic occurs in a population of slightly higher density, its size is equal approximately to twice the excess.

It will be seen that the more complex expression A now replaces the simpler fraction κ/l. In fact, when the rates are constant

$$A = \int_0^\infty \kappa e^{-\int_0^\theta l da} d\theta = \kappa \int_0^\infty e^{-l\theta} d\theta = \frac{\kappa}{l}.$$

Reverting to equation (20) it is clear that p can never be equal to unity, as long as N is finite, so that an epidemic can never affect all the susceptible members of a limited population. Of course it has to be recognised that when the population has been reduced to small numbers the equations here given do not strictly hold.

It may also be pointed out that the population density $N_0 = 1/A$ is only a threshold density with respect to initial importations of cases which have just been infected. That is to say the cases present at the commencement of the epidemic are assumed to be of the type $v_{0,0}$, and none are of the types $v_{0,1}, v_{0,2} \ldots v_{0,r}$. It is this limitation which renders it impossible in the general case to identify the threshold population with the number who are still unaffected at the instant when the epidemic reaches its maximum, since at that instant many cases will certainly be not just commencing but will be of the type $v_{0,r}$, and so they cannot be treated as equivalent to those which we have assumed to have been originally introduced. Nevertheless there seems little doubt that by analogy with the simpler case in which the rates were constants, the

point at which the epidemic reaches its maximum will, in general, correspond approximately with the point at which the remaining unaffected population has been reduced to the threshold value.

Another point of interest arising from equation (20) is in relation to variations in the infectivity rate. It will be seen that the effect of increasing the infectivity from ϕ_θ to $\alpha\phi_\theta$ is to increase A to αA, and consequently the threshold value N_0 is reduced to N_0/α.

Let $\alpha = 1 + \beta$, where β is very small, so that β is the fractional increase in the infectivity.

The new threshold is now $\dfrac{N_0}{1+\beta} = N_0 - \beta N_0$. Consequently the excess being now βN_0, an epidemic of the size $2\beta N_0$ is to be expected. Thus a small increase in the infectivity rate may cause a very marked epidemic in a population which would otherwise be free from epidemic, provided that the population was previously at its threshold value. On the other hand, if the actual density was below the threshold, no epidemic could occur until the infectivity had been increased to such a degree as to make the threshold value less than the actual density.

(12) It is not difficult to extend these results to such diseases as malaria or plague, in which transmission is through an intermediate host. In this case using dashed letters for symbols referring to the intermediate host we have

$$\left.\begin{aligned} \frac{d\log x}{dt} &= \int_0^t A'_\theta v'_{t-\theta}\,d\theta + A'_t y'_0 \\ \text{and} \qquad \frac{d\log x'}{dt} &= \int_0^t A_\theta v_{t-\theta}\,d\theta + A_t y_0 \end{aligned}\right\}, \tag{34}$$

whence

$$\left.\begin{aligned} -\log\frac{1-p}{1-\frac{y_0}{N}} &= A'p'N' \\ \text{and} \qquad -\log\frac{1-p'}{1-\frac{y_0'}{N'}} &= Ap N \end{aligned}\right\}. \tag{35}$$

Neglecting y_0/N and y_0'/N' as before we have to a first approximation

$$p\left(1+\frac{p}{2}\right)p'\left(1+\frac{p'}{2}\right) = AA'pp'NN',$$

thus

$$\frac{p}{2}+\frac{p'}{2} = AA'NN' - 1, \tag{36}$$

As p and p' are always positive where there is an epidemic, $AA'NN'$ must be greater than 1, or a true epidemic can occur only when $AA'NN'$ is greater than unity. We thus see that there is no threshold in the sense used in the previous paragraph for either man or the intermediate host separately, but that there exists what may be called a threshold product $1/AA'$, and this must be exceeded by the product NN' in order that an epidemic may occur.

We shall now suppose that the value of $N' = N_0'$, and that $N = \frac{1}{AA'N_0'} + n$ where n is not very great compared with $1/AA'N_0'$, thus $N = N_0 + n$.

We observe that if the value N had been N_0, the situation would be such that no epidemic could arise. In fact, the product NN' would have been at its threshold value. If, however, N exceeds this value N_0 by an amount n, and if we regard N_0 as remaining fixed, then under this condition N_0 corresponds to a threshold value in the former sense, and we are considering the case in which this threshold value is exceeded by n.

Eliminating p' from the above equations we have to a first approximation

$$p = \frac{2n}{N_0} \frac{A'N_0'}{1 + A'N_0'}. \tag{37}$$

Three cases may be considered :

(1) When N_0' is very small, $p = 0$, and a true epidemic will not occur.

(2) When $N_0' = 1/A'$, $pN_0 = n$.

The size of the epidemic is here exactly equal to the excess and the result of the epidemic is to reduce the population to its threshold value.

(3) When N_0' is very great, $pN_0 = 2n$, or to double the excess.

In this case the size of the epidemic is the same as in the simple case previously considered. That this should be so is apparent, when we consider that the assumption that N_0' is very great, is equivalent to the assumption that the intermediate host is so plentiful that we are dealing with a condition which is practically identical with contact infection.

Further reverting to equation (36) and multiplying both sides by N_0N_0' we have

$$N_0'pN_0 + N_0p'N_0' = 2N_0N_0'(AA'NN' - 1).$$

We choose

$$N_0N_0' = 1/AA' = \pi_0,$$

where π_0 is what we have called above the threshold product. That is to say, when the populations are simultaneously N_0 and N_0' there will be no epidemic. Then

$$\begin{aligned} N_0'pN_0 + N_0p'N_0' &= 2(NN' - N_0N_0') \\ &= 2(\pi - \pi_0), \end{aligned}$$

where π is equal NN′, and we suppose that π is greater than π_0. Now let $\bar{N}$ and $\bar{N}'$ be the populations after the epidemic has terminated, and let $\bar{\pi} = \bar{N}\bar{N}'$.

Then

$$\begin{aligned}\pi - \bar{\pi} &= NN' - (N - \Delta N)(N' - \Delta N'),\\ &= N\Delta N' + N'\Delta N - \Delta N \Delta N',\\ &= Np'N' + N'pN - pNp'N',\\ &= NN'(p + p' - pp'),\\ &= N_0N_0'(p + p' - pp') + (NN' - N_0N_0')(p + p' - pp').\end{aligned}$$

If the excess of population is small so that $NN' - N_0N_0'$ is small as compared with N_0N_0', we can neglect the second term. Further, pp' can be neglected as compared with p or p', and therefore

$$\pi - \bar{\pi} = N_0N_0'(p + p') = 2(\pi - \pi_0). \tag{38}$$

That is to say, the difference between the values of the product of populations before and after the epidemic is twice the excess of the product before the epidemic over the threshold product. This equation is exactly analogous to equation (33). Somewhat similar results have been previously obtained by one of us (McKendrick, 1912) in an analogous but slightly different problem.

(13) These results account in some measure for the frequency of occurrence of epidemics in populations whose density has been increased by the importation of unaffected individuals. They also emphasise the role played by contagious epidemics in the regulation of population densities. It is quite possible that in many regions of the world the actual density of a population may not be widely different from the threshold density with regard to some dominant contagious disease. Any increase above this threshold value would lead to a state of risk, and of instability. The longer the epidemic is withheld the greater will be the catastrophe, provided that the population continues to increase, and the threshold density remains unchanged. Such a prolonged delay may lead to almost complete extinction of the population. Similar results, though of a somewhat more complicated form, hold for epidemics transmitted through an intermediate host. In this case, in place of the threshold density we have to consider the threshold product.

Summary.

1. A mathematical investigation has been made of the progress of an epidemic in a homogeneous population. It has been assumed that complete immunity is conferred by a single attack, and that an individual is not infective at the

moment at which he receives infection. With these reservations the problem has been investigated in its most general aspects, and the following conclusions have been arrived at.

2. In general a threshold density of population is found to exist, which depends upon the infectivity, recovery and death rates peculiar to the epidemic. No epidemic can occur if the population density is below this threshold value.

3. Small increases of the infectivity rate may lead to large epidemics; also, if the population density slightly exceeds its threshold value the effect of an epidemic will be to reduce the density as far below the threshold value as initially it was above it.

4. An epidemic, in general, comes to an end, before the susceptible population has been exhausted.

5. Similar results are indicated for the case in which transmission is through an intermediate host.

REFERENCES.

Fock, ' Math. Zeit.,' vol. 21, p. 161 (1924).
McKendrick, ' Paludism,' No. 4, p. 54 (1912).
McKendrick, ' Proc. Edin. Math. Soc.,' vol. 44 (1925–26).
Ross and Hudson, ' Roy. Soc. Proc.,' A, vols. 92 and 93 (1915–17).

INDEX TO VOL. CXV. (A)

Adsorption, experimental test of dipole theory (Palmer), 227.
Alloy crystals, tensile tests (Elam), 133, 148, 167.
Alloys, iron-manganese, thermal changes (Hadfield), 120.
Aluminium, conductivity of a single crystal (Griffiths), 236.
Appleton (E. V.) and Ratcliffe (J. A.) On the Nature of Wireless Signal Variations, 291, 305.
Askey (P. J.) See Hinshelwood and Askey.
Astbury (W. T.) A Simple Radioactive Method for the Photographic Measurement of the Integrated Intensity of X-Ray Spectra, 640.
Aston (F. W.) A New Mass-Spectrograph and the Whole Number Rule, 487.
Auroral line, green, wave-length (McLennan and McLeod), 515.

Bakerian lecture (Aston), 487.
Bentivoglio (M.) Rate of Growth of Crystals in Different Directions, 59.
Benzene, latent heat of vaporization (Sutcliffe, Lay and Prichard), 88.
Bone (W. A.) and Newitt (D. M.) Gaseous Combustion at High Pressures—VII, 41.
Bradley (A. J.) and Thewlis (J.) The Crystal Structure of α-Manganese, 456.
Brotherton (M.) See Richardson and Brotherton.
Brownian movement of a galvanometer coil (Ornstein and others), 391.
Burger (H. C.) See Ornstein and others.

Callendar (L. H.) The Influence of Boundary Films on Corrosive Action, 349.
Carbon dioxide, dissociation at high temperatures (Fenning and Tizard), 318.
Caress (A.) and Ridcal (E. K.) The Combination of Nitrogen and Hydrogen Activated by Electrons, 684.
Cathode rays, slow, absorption in various gases (Lehmann), 624
Cathode rays, slow, total ionisation due to absorption (Lehmann and Osgood), 609.
Chapman (S.) On Certain Average Characteristics of World-Wide Magnetic Disturbance, 242.
Charcoal, expansion on sorption of carbon dioxide (Meehan), 199.
Clarkson (W.) See Ornstein and others.
Constable (F. H.) The Cause of the Colours shown during the Oxidation of Metallic Copper, 570.
Cook (W. R.) See Lennard-Jones and Cook.
Copper oxidation colours (Constable), 570.
Corrosive Action, influence of boundary films (Callendar), 349.
Crystal, single aluminium, conductivity (Griffiths), 236.
Crystal structure of α-manganese (Bradley and Thewlis), 456.
Crystalline nitrates and carbonates, magnetic anisotropy (Krishnan and Raman), 549.
Crystals, rate of growth in different directions (Bentivoglio), 59.

Darwin (C. G.) The Zeeman Effect and Spherical Harmonics, 1.
Dennison (D. M.) A Note on the Specific Heat of the Hydrogen Molecule, 483.
Dilatometer, new differential (Smith), 554.
Dimethyl ether, gaseous kinetics of decomposition (Hinshelwood and Askey), 215.

Elam (C. F.) Tensile Tests on Alloy Crystals, 133, 148, 167.
Electric moment of the sulphur complex (Taylor and Rideal), 589.
Electron emission under influence of chemical action (Richardson and Brotherton), 20.
Electrons, combination of nitrogen and hydrogen activated by (Caress and Rideal), 684.
Epidemics, mathematical theory (Kermack and McKendrick), 700.

Fenning (R. W.) and Tizard (H. T.) The Dissociation of Carbon Dioxide at High Temperatures, 318.
Fisher (J. W.) See Flint and Fisher.
Flint (H. T.) and Fisher (J. W.) A Contribution to Modern Ideas on the Quantum Theory, 208.

Galvanometer coil, brownian movement (Ornstein and others), 391.
Gaseous combustion at high pressures—VII (Bone and Newitt), 41.
Gaseous mixture, equation of state (Lennard-Jones and Cook), 334.
Griffiths (E.) The Thermal and Electrical Conductivity of a Single Crystal of Aluminium, 236.

Hadfield (Sir Robert) Thermal Changes in Iron-Manganese Alloys, low in Carbon, 120.
Havelock (T. H.) The Method of Images in Some Problems of Surface Waves, 268.
Hinshelwood (C. N.) and Askey (P. J.) Homogeneous Reactions involving Complex Molecules—The Kinetics of the Decomposition of Gaseous Dimethyl Ether, 215.
Homogeneous reactions—decomposition of gaseous dimethyl ether (Hinshelwood and Askey), 215.
Horrocks (H.) Meteorological Perturbations of Tides and Currents in an Unlimited Channel Rotating with the Earth, 170; Generalised Sturm-Liouville Expansions, 184.
Hydrogen molecule, specific heat (Dennison), 483.

Images, method of, in problems of surface waves (Havelock), 268.
Iron-manganese alloys, thermal changes (Hadfield), 120.

Jack (D.) The Band Spectrum of Water Vapour, 373.

K series emission by K ionised atoms (Martin), 420.
Kapitza (P.) Further Developments of the Method of obtaining Strong Magnetic Fields, 658.
Kermack (W. O.) and McKendrick (A. G.) A Contribution to the Mathematical Theory of Epidemics, 700.
Krishnan (K. S.) and Raman (C. V.) The Magnetic Anisotropy of Crystalline Nitrates and Carbonates, 549.

Lay (F. C.) See Sutcliffe, Lay and Prichard.
Lehmann (J. F.) The Absorption of Slow Cathode Rays in Various Gases, 624.
Lehmann (J. F.) and Osgood (T. H.) The Total Ionisation due to the Absorption in Air of Slow Cathode Rays, 609.
Lennard-Jones (J. E.) and Cook (W. R.) The Equation of State of a Gaseous Mixture, 334.

McKendrick (A. G.) See Kermack and McKendrick.
McLennan (J. C.) and McLeod (J. H.) On the Wave-Length of the Green Auroral Line in the Oxygen Spectrum, 515.
McLeod (J. H.) See McLennan and McLeod.
Magnetic anisotropy of crystalline nitrates and carbonates (Krishnan and Raman), 549.

Magnetic disturbance, world wide, average characteristics (Chapman), 242.
Magnetic fields, strong, method of obtaining (Kapitza), 658.
Manganese, alpha, crystal structure (Bradley and Thewlis), 456.
Manganese-iron alloys, thermal changes (Hadfield), 120.
Martin (L. H.) The Efficiency of K Series Emission by K Ionised Atoms, 420.
Mass-spectrograph and the whole number rule (Aston), 487.
Meehan (F. T.) The Expansion of Charcoal on Sorption of Carbon Dioxide, 199.

Newitt (D. M.) See Bone and Newitt.
Nitrogen and hydrogen, combination activated by electrons (Caress and Rideal), 684.

Ornstein (L. S.), Burger (H. C.), Taylor (J.) and Clarkson (W.) The Brownian Movement of a Galvanometer Coil, 391.
Osgood (T. H.) See Lehmann and Osgood.
Oxidation of metallic copper, cause of the colours (Constable), 570.

Palmer (W. G.) An Experimental Test of the Dipole Theory of Adsorption, 227.
Paris (E. T.) On the Reflexion of Sound from a Porous Surface, 407.
Photo-electric experiments with liquid alloys (Richardson and Brotherton), 20.
Polytope in space of n-dimensions, relations of angle-sums and volume (Somerville), 103.
Prichard (W. Ll.) See Sutcliffe, Lay and Prichard.

Quantum theory, contribution to (Flint and Fisher), 208.

Raman (C. V.) See Krishnan and Raman.
Ratcliffe (J. A.) See Appleton and Ratcliffe.
Richardson (O. W.) The Hydrogen Band Spectrum: New Band Systems in the Violet, 528.
Richardson (O. W.) and Brotherton (M.) Electron Emission under the Influence of Chemical Action at Higher Gas Pressures and some Photo-electric Experiments with Liquid Alloys, 20.
Richardson (O. W.) and Robertson (F. S.) The Emission of Soft X-Rays by different Elements, 280.
Rideal (E. K.) See Caress and Rideal, and Taylor and Rideal.
Robertson (F. S.) See Richardson and Robertson.
Rusting of surfaces in contact (Tomlinson), 472.

Smith (C. J.) A New Differential Dilatometer for the Determination of Volume Changes during Solidification, 554.
Somerville (D. M. Y.) The Relations connecting the Angle-Sums and Volume of a Polytope in Space of n-Dimensions, 103.
Sorption of carbon dioxide and expansion of charcoal (Meehan), 199.
Sound, reflexion from a porous surface (Paris), 407.
Spectrum, band, of water vapour (Jack), 373.
Spectrum, hydrogen, new band systems in the violet (Richardson), 528.
Spectrum, oxygen, wave-length of green auroral line (McLennan and McLeod), 515.
Spherical harmonics and Zeeman effect (Darwin), 1.
Sturm-Liouville expansions, generalised (Horrocks), 184.
Sulphur complex, electric moment (Taylor and Rideal), 589.
Sutcliffe (J. A.), Lay (F. C.) and Prichard (W. Ll.) The Latent Heat of Vaporization of Benzene at Temperatures above the Boiling Point, 88.

Taylor (A. M.) and Rideal (E. K.) On the Electric Moment of the Sulphur Complex, 589.
Taylor (J.) See Ornstein and others.
Tensile tests on alloy crystals (Elam), 133, 148, 167.
Thewlis (J.) See Bradley and Thewlis.
Thunderclouds and showers, currents carried by point discharges (Wormell), 443.
Tides and currents, meteorological perturbations (Horrocks), 170.
Tizard (H. T.) See Fenning and Tizard.
Tomlinson (G. A.) The Rusting of Surfaces in Contact, 472.

Water vapour, band spectrum (Jack), 373.
Waves, surface, method of images applied (Havelock), 268.
Wireless signal variations (Appleton and Ratcliffe), 291, 305.
Wormell (T. W.) Currents carried by Point-Discharges beneath Thunderclouds and Showers, 443.

X-Ray Spectra, radioactive method for photographic measurement of integrated intensity (Astbury), 640.
X-Rays, soft, emission by elements (Richardson and Robertson), 280.

Zeeman effect and spherical harmonics (Darwin), 1.

END OF THE ONE HUNDRED AND FIFTEENTH VOLUME (SERIES A).